Mental Effort and Sustained Perceptual-Motor Performance

By

Stephen Henry Fairclough

Thesis

Submitted in partial fulfilment of the requirements for the award of

Degree of Doctor of Philosophy of Loughborough University

2nd February 2001-02-14

ACKNOWLEDGEMENTS

"OK brain, I don't like you and you don't like me, but let's just do this thing and then I can get back to killing you with beer"

H. J. Simpson

As I come to the end of this very long and winding road, I would like to express sincere gratitude to those who helped me along the way. First of all, I'd like to thank my supervisor Andrew Shepherd for his support and astute critical input, which more than compensated for his deeply prejudicial views on Wigan R.L.F.C. I'd like to thank my colleagues at the HUSAT Research Institute who supported the experimental work described in this document, particularly, Jayshree Lakha (who scheduled the subjects for experiments 3 and 4 despite a horrendous matched-subjects design), Katherine Hack (who did the subject scheduling for experiment 1), Rob Graham (my co-worker on experiment 3), Steve Hirst (who endured my ramblings on various topics with great sympathy) and Peter Marsh (who wrote data analysis software for experiment 3).

I'd like to acknowledge the support of those who went before me - Drs. Terry Lansdown and Peter Burns, particularly during those long and intensive lunches at the sadly-departed Bombay Mix. Moving further afield, I'd like to say a big thank you to my Dutch friends who have provided me with technical support and personal encouragement through the years: in particular, Dick De Waard and Karel Brookhuis from COV at the University of Groningen, Willem Verwey and Wim Van Winsum (who provided the necessary software for experiment 4). I'd also like to express thanks to those researchers based in the U.S. who were good enough to provide me with paper reprints and encouragement: Gerry Matthews, Paula Desmond, Peter Hancock, Joel Warm and Mark Scerbo.

Moving away from fellow professionals, I'd like to acknowledge various sources of personal support ("how's it going?") and interest ("haven't you finished it yet?"): my parents, Margaret and Terry, my brother, Stuart; our Nottingham inner-circle (Sean, Fiona, Ian and my in-laws, Pam and Mick), the fabled Watkins of Blaenau Ffestiniog, Brian and Andrea (Wigan's leading aquarists) and my old student buddy Ken (and Cath).

I've saved my biggest vote of thanks as the last. I'd like to express deep and heartfelt thanks to my wife Steph, who has lived with this "thing" for as long as I have and supported me at every twist and turn.

ABSTRACT

The thesis is concerned with the concept of mental effort and sustained perceptual-motor performance. It is proposed that mental effort represents a finite resource which is capable of modulating those cognitive operations underlying perceptual-motor performance. In addition, it is hypothesised that mental effort investment is associated with a number of energetical costs. This finite quality of effort is used to justify a self-regulatory mechanism designed to ensure that effort is deployed in a rational fashion.

A model of mental effort regulation is proposed based on a process of self-appraisal. This process is informed via feedback from external sources (concerned with performance effectiveness) and internal symptoms (concerned with energetical state and personal discomfort). Both sources of feedback are amalgamated into a cognitive-energetic appraisal, which represents the basis for the formulation of effort policy, i.e. whether to increase, decrease or sustain current level of mental effort.

This model is explored via four experimental studies of sustained perceptual-motor performance. An initial laboratory study investigates the influence of individual differences and sleep deprivation on effort regulation. This study revealed that low performance capacity is associated with an increase of mental effort investment. The following two studies are concerned with driving behaviour. An experiment contrasting the influence of sleep deprivation and alcohol on driving performance illustrates the influence of external and internal feedback on effort regulation. The following experiment investigates this issue in detail by exposing drivers to objective performance feedback during a sustained 'journey'. A final laboratory study provides evidence that mental effort is a finite commodity and illustrates a number of alternative effort policies in action.

These findings and their implications for the model of mental effort regulation are discussed.

1 INTRODUCTION

1.1 Definitions and hypotheses

This thesis is based on two related propositions: (a) that mental effort exerts an influence on performance quality and energetical state, and (b) the regulation of mental effort is determined via a dual assessment of external feedback (related to overt performance quality) and internal feedback (related to energetical state). The purpose of the thesis is to investigate the conceptual utility and empirical validity of both propositions within a model of mental effort regulation.

This model is proposed and investigated within the context of sustained perceptual-motor performance. This task scenario has been selected to focus the scope of the current document onto those sustained activities in the laboratory and in the field which permit a degree of self-regulation.

The concept of mental effort has growing significance for work psychology and ergonomics. In the course of the current century, from the industrial revolution onwards, Western working life has been characterised by a shift of emphasis from manual to mental activities. This trend is maintained and accelerated by contemporary developments in information technology. For instance, a pan-European survey reported by Paoli (1997) conducted from 1995 to 1997 indicated that "time pressure" in the workplace had increased from 35 to 42%. However, an upper limit on the efficiency and safety (and therefore, the benefits) of current and future technological systems is determined by the inherent limitations of the human information processing system.

It is proposed that mental effort may exert a crucial influence on human performance capability, both in the sense of task demands and maximum working duration, i.e. how much information is too much information? How long can individuals sustain performance on a given task? In this respect, mental effort has a pivotal role if good human factors (with respect to both technology and working practices) is defined by a reciprocal process of performance optimisation between human and machine.

It may be argued that mental effort is a pertinent and significant construct for the study of human performance. However the concept of mental effort does not enjoy a high profile in either ergonomics research or the annals of psychological literature. There are several possible reasons for this relative neglect. In the first instance, the concept of effort has traditionally been associated with physical activity. This confusion stems from early fatigue research where physical effort was defined in concrete terms using measures of muscular metabolism (Ryan, 1953). As we shall see, the quantification of mental effort is not so straightforward. The current thesis is not concerned with physical effort or physical work activity. Therefore, future usage of both terms "effort" and "mental effort" are intended to be synonymous.

A second encumbrance originates from particular problems of definition associated with effort. The absence of consistent definition has rendered effort little more than a conceptual dumping ground for a variety of vague, motivational variables over the years. It is argued that such difficulties owe much to a problem of semantic transition, when words are extrapolated from colloquial usage into the realm of theory. Specifically, the word 'effort' has semantic ties to a number of everyday verbs, e.g. to scrutinise, to focus, to concentrate - each of which contribute to a colloquial definition. This observation is not intended as criticism in itself - linkage between a scientific term and a colloquial term does not imply that the former is necessarily flawed by association with the latter. However, there is a requirement to make a clear distinction between scientific and colloquial definitions of the term. A number of "common-sense" assumptions are inherent in colloquial usage, rooted in shared understanding, which should not infiltrate scientific definition in the absence of an explicit theoretical or empirical justification.

This impact of colloquial "noise" in scientific writing may impede the development of shared definition and understanding within the research community. A more insidious influence concerns the generation of post-hoc hypotheses based on colloquial assumptions. It is argued that effort is particularly vulnerable in this respect. If we feel that we understand what it means experientially to "concentrate" or "intensify effort" – this may inhibit the development of scientific explanations to describe the same phenomenon. Alternatively, a researcher finding a significant and unanticipated improvement of performance may cite "increased effort" as a post-hoc explanation without the requirement for further explanation. All of which conspire to discourage detailed explanation and investigation.

A third inhibitory factor concerns the location of mental effort within psychological traditions and sub-disciplines. Hockey, Coles, & Gaillard (1986) illustrated that effort had an equidistant relationship with both "wet" biological and "dry" cognitive or computation traditions in psychology. In other words, mental effort is presumed to modulate attentional processing within the cognitive system.

However, the manifestation of mental effort was frequently registered as psychophysiological activity in the central nervous system, an area allied to biological tradition in psychology (Mulder & Van Der Meulen, 1973). The analysis produced by Hockey, Coles, & Gaillard (1986) stipulated the existence of a schism between both traditions and formulated the concept of "energetics" to bridge the gap. Therefore, the concept of mental effort may have been developed inconsistently due to the biological and computational bias of individual researchers. It is only recently that a unified, cognitive-energetical view of mental effort has been proposed (Hockey, 1993; Hockey, 1997).

Despite these obstacles, mental effort has endured as a psychological concept. It has been proposed that the existence of an effort concept is based upon three central hypotheses (Pashler, 1998):

- The investment of mental effort will generally result in an improvement of task performance.
- The investment of mental effort is associated with certain psychological "costs."
- The regulation of mental effort is under volitional control.

Therefore, mental effort does not exist as a unitary construct, but rather as a connected group of related hypotheses that are rarely investigated simultaneously. The following paragraphs will expand and analyse each hypothesis in detail.

The first hypothesis states that mental effort will improve the <u>quality of performance</u>. There are several aspects to this assertion. In the first instance, it is assumed that an investment of mental effort is associated with an intensification of psychological activity, which brings about an increase of performance quality. In other words, the cognitive system may work at a higher rate or to a higher level of precision when mental effort is increased. Wickens (1986b) proposed a gain function in order to describe this intensification of cognitive activity. Within his analysis, a vector equation was used to describe a performance output (X) resulting from a series of information processing operations (A, B, C):

$$\mathbf{X = [A, B, C],}$$

In Wickens' analysis, each of the three information processing operations (A, B, C) were associated with a gain function. Therefore, the vector equation was altered as follows:

$$\mathbf{X = [Ga, Gb, Gc] * [A, B, C]}$$

Wickens (1986b) proposed that gain could exert an influence on information processing in two ways. In the first instance, a cognitive process was modulated with respect to the signal-to-noise ratio. This type of modulation may influence the sensitivity and hence the evaluative quality of a given cognitive component, e.g. to isolate a target from noise during visual search. Secondly, gain could affect those qualitative aspects of information processing such as an alteration in terms of response bias, e.g. to recognise red shapes rather than blue shapes in a visual array. Both forms of modulation represent possible mechanisms by which mental effort could influence the quality of performance.

Regardless of the specific mechanism mediating effort and performance, an individual will always invest mental effort with the express purpose of performance improvement or protection. This distinction was characterised by Mulder (1986) as effort investment in the service of task demands (i.e. task effort) versus effort investment to buttress performance in the presence of physical and environmental stressors such as drugs, sleep deprivation, noise, high temperature etc. In the case of the former, mental effort is invested as a response to a perceived increase of task demand, e.g. complexity, pacing etc. The investment of state effort is primarily concerned with performance reinforcement under conditions of duress. Mulder (1986) suggested that both exist as distinct modes of mental effort investment. The current thesis argues that task- and state-related effort represent different instances of the same underlying mechanism (this statement will be expanded at a later point).

One problem associated with the first hypothesis is an obvious disregard for the possibility that increased mental effort may not always achieve a desired improvement of performance. This scenario will certainly arise if an individual has insufficient expertise or inadequate sensory capabilities. In addition, the investment of mental effort may not improve performance of elementary task activities. For example, consider how the investment of mental effort may affect the performance of a simple mathematical task, e.g. 2 + 2. In certain instances, the quality of performance may be insensitive to varying levels of mental effort investment. This possibility was articulated by Norman & Bobrow (1975) as *data-limited* performance.

Eysenck (1982) and Schonpflug (1985) postulated a complimentary argument - that the quality of performance could be expressed in terms of *efficiency*. This conceptualisation represented a ratio relationship between cognitive intensity or effort and overt performance. Within this framework (extended in Eysenck & Calvo (1992)), performance matrices are characterised as overt manifestations of internal cognitive operations. If a high standard of performance is possible with minimum effort investment, this represents the highest level of efficiency. In other cases, performance effectiveness may only be achieved "at cost," i.e. via a corresponding investment of mental effort. A matrix representing the relationship between mental effort and performance effectiveness is shown below in Figure 1. The descending arrow from left to right is used to represent declining efficiency as high performance is attained at higher levels of "cost" in terms of mental effort. The arrow descending from right to left illustrates how effectiveness may only be achieved with a corresponding investment of mental effort.

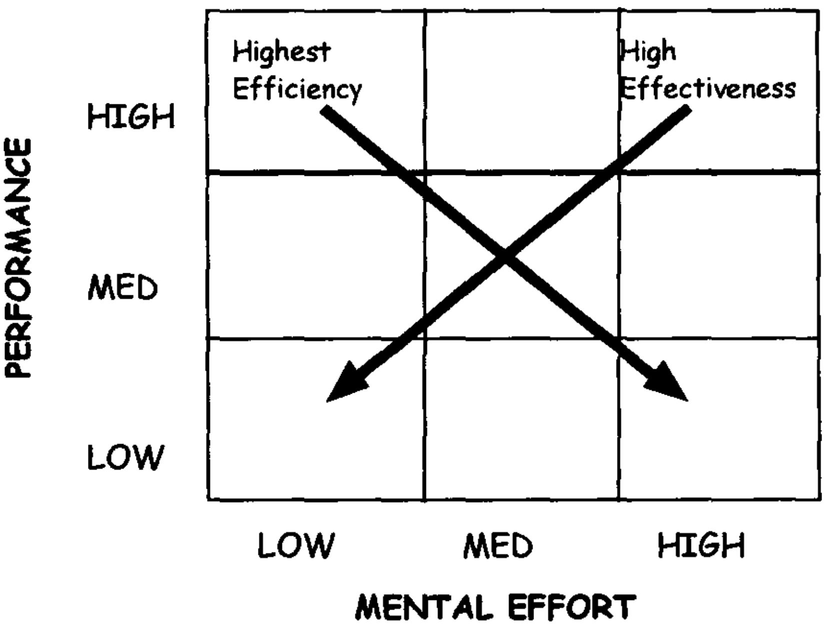


Figure 1. The relationship between performance effectiveness and mental effort based on the analysis of Eysenck & Calvo (1992).

The relationship shown in Figure 1 exemplifies the second hypothesis concerning the <u>costs</u> associated with mental effort investment. These economical metaphors of "costs" and "expenditure" represent a legacy from the model of mental effort described by Kahneman (1973) and other resource theorists (Navon & Gopher, 1979). For Kahneman and others, mental effort was assumed to exist as a <u>finite</u> commodity, present in "reserve" and capable of being in "debt" if overextended (similar to the oxygen debt that characterises muscular exertion).

The costs associated with mental effort are due to the finite nature of effort reserves and may manifest themselves in a number of ways. This topic has been addressed in a series of papers by Hockey (Hockey, 1986b; Hockey, 1993; Hockey, 1997). His argument is that mental effort regulation is highly sensitive to a number of *latent decrements*, such as selective task failure, increased sympathetic nervous system activity and fatigue after-effects. These latent decrements represent various costs associated with mental effort investment, which may be apparent before any degradation of task performance. It should be noted that costs might express themselves via an impact on overt behaviour/strategy or in terms of psychophysiological changes. In addition, rising intensity/frequency of latent decrement may function on the basis of a positive feedback loop, i.e. actively degrading the quality of performance, demanding further effort investment and subsequently acerbating the frequency and magnitude of future latent decrements.

Schonpflug (1986a) characterised the relationship between mental effort and performance quality in terms of "Behaviour Economics." Within his framework, the costs of exertion due to mental effort investment included: (a) reduced physical and mental health, (b) exhaustion of internal biological energy, and (c) exhaustion of external resources such as social assistance and technical aids. According to his view, mental effort must be regulated in accordance with three broad strategies: (1) to save effort whenever possible, (2) to calculate the utility of effort deployment (in terms of expected outcomes and consequences), and (3) to rationalise the distribution of mental effort over time-on-task. The issue of utility is key to an understanding of "Behaviour Economics" and the analysis of effort shown in Figure 1.

A connection between mental effort and voluntary control (i.e. the third hypothesis) was described by Binswanger (1991). According to his Objectivist perspective, the purpose of mental effort was to fuel a "conceptual faculty" thus enabling "mental focus." The development of "focus" implied: logical introspection, the identification of vagueness, a counteraction of emotional bias (an awareness of thinking rather than feeling), a distinction between rationality and rationalisation, and a consideration of a wide range of evidence (i.e. "a mind in full focus places no consideration above the truth" (Binswanger, 1991), p. 164). In a manner similar to those economic notions in the previous paragraph, this perspective places great emphasis on rationality. For instance, if an individual desires a particular outcome or goal, he or she employs the conceptual faculty in order to formulate a plan that is both effective and efficient.

The underlying notion of this hypothesis is the volitional regulation of mental effort. In other words, an individual may decide whether to invest effort into a task. This subject of volition was heavily mooted by late 19th century psychologists and has been prominent by its absence from contemporary research (Baars, 1993; Kimble & Perlmuter, 1970). Despite this low profile, Baars (1992) proposed that the ideomotor theory of voluntary control devised by James (1890) represented a strong foundation for contemporary research. In his original work, James (1890) emphasised the formulation of psychological goals to enable voluntary control. For instance, James (1890) employed the example of man awakening in the morning and attempting to leave a warm bed in a cold room. The process of voluntarily leaving the bed was achieved by overwhelming one "image" or goal (the anticipated cold and discomfort of getting out of bed) with another goal (the obligation to get up and engage in purposive activity). It is presumed that effort investment plays a role in the resolution of antagonistic goals.

These three hypotheses have developed in relative isolation from one another with different emphases on behaviour. The devolution of effort hypotheses may be blamed on the absence of a coherent and enduring theoretical framework. The current document is concerned with mental effort regulation and sustained performance. An attempt to unite all three hypotheses within a single framework of sustained task performance is presented in the following section.

1.2 Cost and the effort economy

The effort economy is a phrase to describe the expenditure of mental effort across time-on-task. This analogy is based on an economic system of supply and demand (Kahneman, 1973; Navon & Gopher, 1979; Schonpflug, 1986a). This conceptualisation rests on an assumption that mental effort exists as a finite commodity that may be present in "reserve." A refreshed individual begins a task with a full "tank of fuel" or reserve, which is expended at a given rate over the task period. A difficult, demanding task will necessitate a higher rate of effort expenditure to guarantee adequate performance relative to an easier task. Therefore, effort expenditure occupies a two-dimensional space between the requisite level necessary to achieve adequate performance and the maximum duration of task activity.

When the finite reserve of effort is exhausted, adequate performance is no longer possible. Therefore, the rate of effort expenditure and the level of effort reserves before task activity will determine the maximum duration of adequate performance. The relationship between demands (requisite expenditure per unit time) and supply (level of reserves minus cumulative expenditure) provides a simplistic basis for the effort economy. A similar argument has been forwarded by Davies & Parasuraman (1982) to explain the performance decrement characteristic of vigilance activities (i.e. a sustained, perceptual vigil).

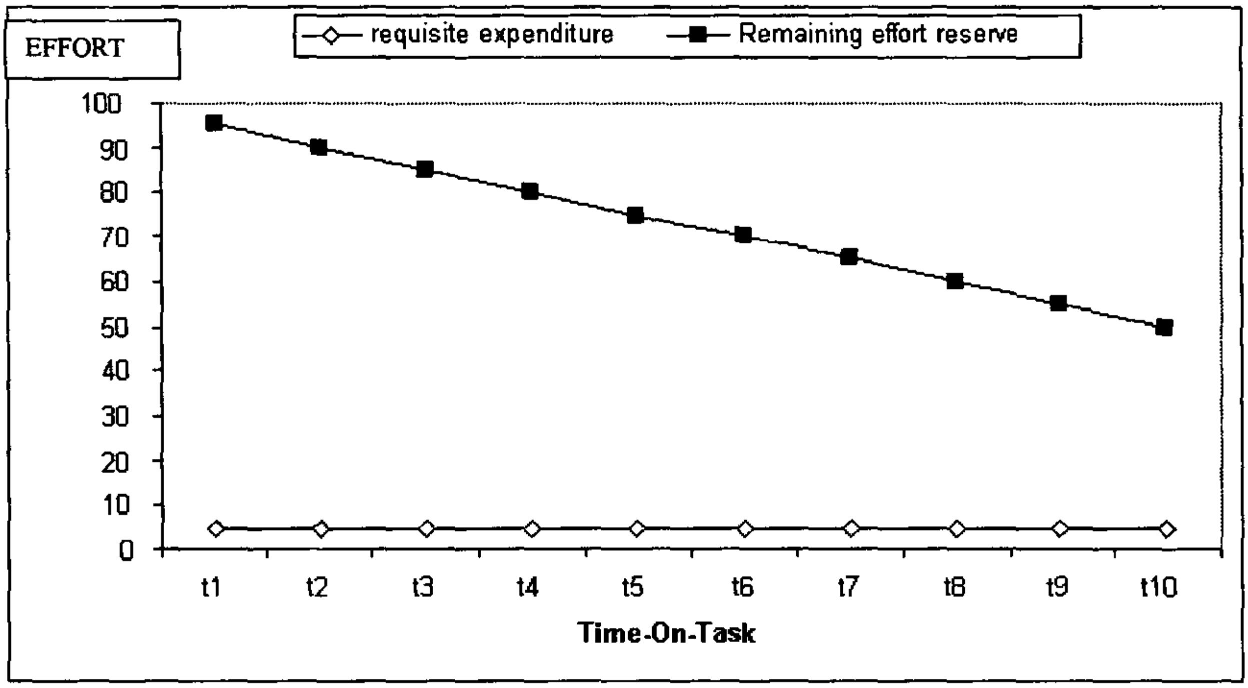
The level of task demand determines the level of expenditure necessary to sustain adequate performance, however, it is not the only factor capable of influencing requisite expenditure. If the individual must perform in the presence of internal stressors (e.g. sleeplessness, drugs) or external stressors (e.g. noise, extreme heat), these energetical factors will inflate task demands and the requisite level of effort expenditure. This is an important point as certain energetical factors such as sleeplessness and fatigue may depress the level of effort reserves available to an individual prior to task activity. Therefore, energetical stressors may be capable of influencing both expenditure and level of reserves available for task performance.

The simplest version of the effort economy is analogous to a fuel-tank metaphor and is illustrated in Figure 2. In both examples, effort investment is determined by three variables: (a) the amount of effort held in reserve, (b) the level of effort expenditure necessary for adequate task performance, and (c) the cumulative decline of effort reserves across time-on-task. In Figure 2a, the requisite level of effort expenditure is modest relative to available reserves. This individual begins the task with 100 effort units in reserve and has only expended half of those reserves at t10.

When task demand is increased as shown in Figure 2b, a different pattern is apparent. This individual must expend effort at 11 units per period and may only sustain a requisite level of effort investment until t8. After this point, the individual is unable to expend sufficient effort to guarantee adequate performance and effort reserves are exhausted by the end of the task.

The scenario illustrated in Figure 2b raises the question – what happens to the performance of those individuals who have exhausted their available effort reserves? Several options are possible. The individual may choose to withdraw from task activity or reach a point of complete physical exhaustion (Hancock & Warm, 1989). Alternatively, an individual may continue to perform at a lower level of effort expenditure or even in the absence of effort expenditure (Hockey, 1993; Hockey, 1997). In this case, two possibilities may emerge: (a) the individual maintains performance albeit at much reduced level or (b) performance collapses to a catastrophic degree.

(a) Low-effort task (requisite effort = 5 units per time period (t))



(b) High-effort task (requisite effort = 11 units per time period (t))

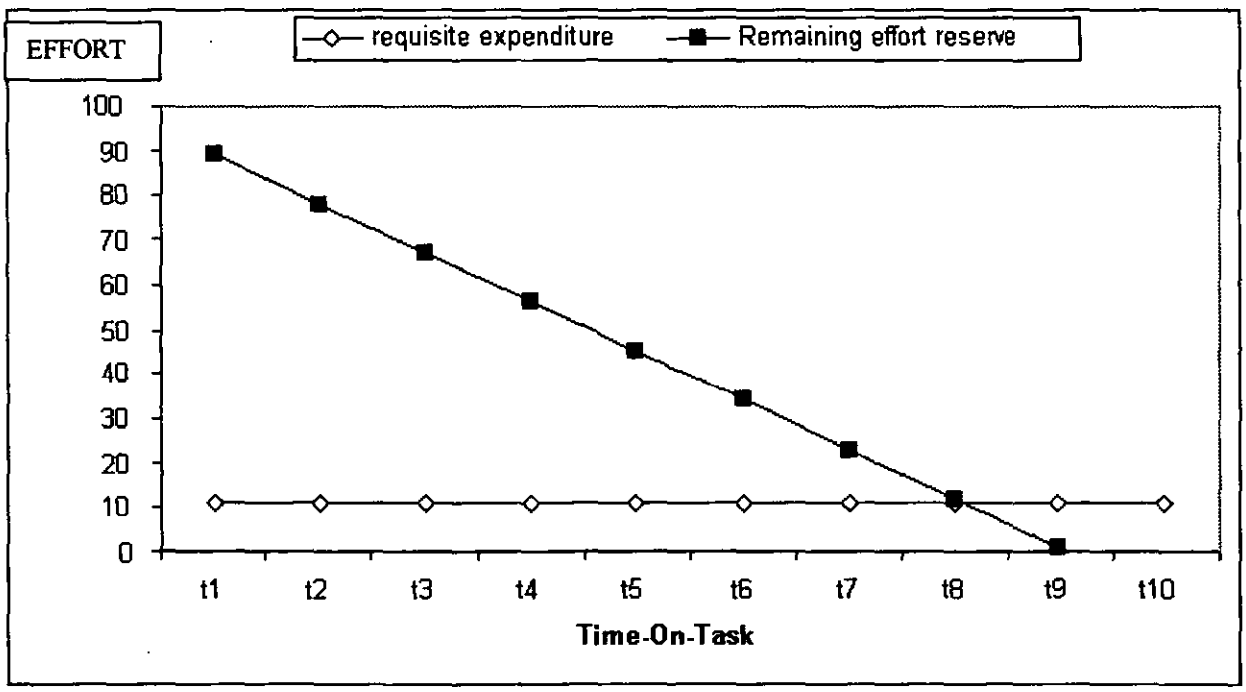


Figure 2. A hypothetical effort economy for sustained investment during (a) a low effort task and (b) a higher effort task.

Given a Hobson's choice of physical collapse, task withdrawal or significantly reduced performance, a state of negative equity within the effort economy is something to be avoided by the human operator. This argument reinforces the desirability of a sensitive and rational mechanism for mental effort regulation, in order to avoid this scenario.

A more 'intelligent' version of the effort economy would permit a proactive type of regulation, wherein effort may be withheld as well as invested in order to avoid a complete exhaustion of reserves. There are two important innovations at the basis of a more advanced model of the effort economy. The first concerns the role of energetical costs associated with continued task performance and effort investment. It has been hypothesised that sustained effort investment is associated with costs such as task stress and fatigue (Hockey, 1993; Hockey, 1997). On this basis, it is postulated that stress, fatigue and related energetical costs may manifest themselves when remaining effort reserves are low. These costs may function as an early-warning signal of impending exhaustion. In addition as stated earlier, energetical costs may render the task more demanding and raise the requisite level of effort expenditure. This mixture of increased discomfort and accelerated costs represents a form of psychological inertia with respect to continued task performance. Furthermore, this inertia may be boosted if the magnitude of costs increase in line with impending exhaustion. If this were the case, the net result of energetical costs would be to produce a rising disinclination to continue task activity (Bartley & Chute, 1947).

The presence of costs may create a chain reaction within the effort economy to reduce the level of expenditure. First, the individual makes a voluntary choice to settle for a lower quality of task performance. This tactic may involve tolerating a higher number of errors or working at a slower pace. The aim of this strategy is to reduce the level of effort expenditure as a means of postponing the exhaustion of remaining effort reserves and ameliorating any discomfort introduced by energetical costs, i.e. effort conservation (Hockey, 1993; Hockey, 1997).

A hypothetical example of this mechanism is presented in Figure 3. This model plots the falling level of effort reserves by requisite and actual levels of effort expenditure, i.e. requisite for acceptable level of performance. The influence of energetical costs on effort expenditure is apparent between time-periods 5 and 6, when the level of requisite expenditure is accelerated with consequent effects on effort reserves. The individual decides to reduce effort and both rates of expenditure dissociate at time-period 7. It is anticipated that the quality of performance is in decline from this point onwards. However, this strategy represents a means by which the individual can reduce their personal discomfort and avoid complete exhaustion of effort reserves (Hockey, 1993; Hockey, 1997).

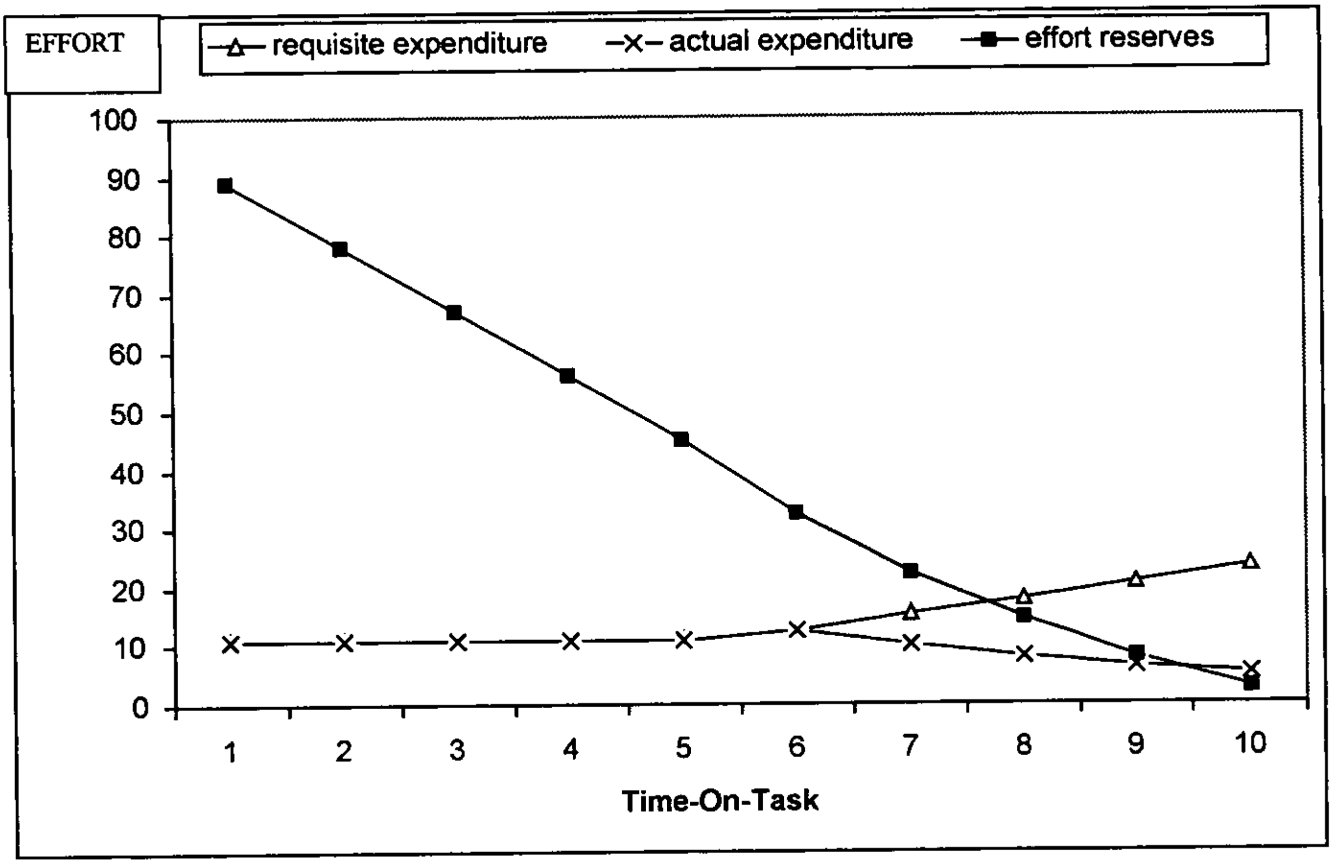


Figure 3. A hypothetical effort economy for a high demand task (11 effort units per time period).

This adaptive mechanism for the management of the effort economy must take account of several important variables. However, it is unlikely that individuals may consult internal indicators for effort reserves and rates of expenditure directly. It is difficult to locate these factors within the experiential world of the individual. It is proposed that critical indicators of the effort economy express themselves indirectly. For example, the level of energetical costs such as stress and/or fatigue may function as surrogate indicators for the level of remaining effort reserves. Similarly, the frequency and magnitude of errors from the primary task may represent the gap (if any) between requisite and actual levels of effort expenditure. These indirect expressions of the effort economy must be monitored and evaluated by a regulatory mechanism.

1.3 The Regulation of Mental Effort

The effort economy is fundamental to the proposed link between volition and cost underlying the regulation of mental effort. This dynamic between the need to performance and the desire to rest rests on the finite characteristic of mental effort, which may not be invested indefinitely with respect to time-on-task. This framework is intended to unite elements of James' ideomotor model (i.e. goal conflict) and those economic models of human information processing (Kahneman, 1973; Navon & Gopher, 1979).

The finite nature of mental effort may be translated into a requirement for a psychological form of 'book-keeping,' wherein investments are registered and evaluated, and debt must be avoided. When an individual faces a potential source of mental effort input, a decision must be made (consciously or otherwise) to invest or to conserve mental effort (Hockey, 1993; Hockey, 1997). Given that effort was assumed to be finite, it is important that this decision is made on a rational basis, taking account of the desires and the capability of the individual. This decision is termed the formulation of effort policy.

The formulation of effort policy is based on feedback with respect to both the quality of ongoing performance and energetical costs. This formulation is determined by external feedback from task performance, i.e. how well am I performing? Similarly, energetical costs may express themselves via internal feedback concerning wellbeing and personal discomfort, i.e. how am I feeling?

It is hypothesised that both performance and discomfort feedback are associated with standards based on past experience. When individuals formulate an effort policy, they have certain expectations concerning the level of performance quality and magnitude of energetical costs associated with a given level of performance. If these expectations are fulfilled, then performance is achieved "at cost." If performance is achieved at a lower psychological cost than expected, this may be termed efficient performance (Eysenck & Calvo, 1992; Schonpflug, 1985). On the other hand, if performance evoked a higher level of psychological cost than expected, the individual may perceive performance to be inefficient.

The concept of behavioural efficiency is crucial to the type of 'book-keeping' that underlies mental effort regulation. Efficiency represents a unification of external and internal sources of feedback into a cognitive-energetical assessment. The assessment of efficiency is the basis of the effort policy. We may choose to invest or to conserve effort in response to inefficiency. It is hypothesised that these effort strategies may be enforced to different quantitative degrees, e.g. high investment, low investment etc.

The three hypotheses of mental effort may be amalgamated into a model of effort regulation, it is assumed: (a) that mental effort influences the quality of performance, and (b) this capacity for modulation is finite and effort is not available in an infinite quantity. In addition, sustained investment of mental effort may be associated with a number of "costs" which may impact on performance quality and energetical status. Therefore, (c) a model of effort regulation is proposed wherein the influence of effort deployment on both performance quality and energetical status is monitored and assessed. This process of monitoring and regulation drives the formulation of effort policy (i.e. a decision whether or not to invest mental effort), which in turn, underlies the volitional nature of effort investment.

The regulation of mental effort is assumed to be a fluid and highly adaptive process. Successful regulation is based on inconstant directives, some of which may even be mutually exclusive, i.e. goal (a) to sustain performance and goal (b) to minimise symptoms of task stress. In addition, as effort is finite in nature, the range of effort strategies available to the individual may be limited in scope on certain occasions.

The model of mental effort regulation adopted by the thesis is simplistic. Feedback is received from dual sources of external and internal feedback. External feedback is concerned with overt performance and the perception of performance efficacy. Internal feedback is perceived along a energetical dimension, concerning the appraisal of Central Nervous System (CNS) functioning and physiological symptoms combined with the self-assessment of affective changes such as mood. The feedback from cognitive task performance and energetical state are combined to produce a cognitive-energetical assessment of behavioural efficiency. This assessment is the basis for the formulation of effort policy.

A schematic representation of this model is shown below in Figure 4. This model will be extended and described in greater detail via a literature review in Chapter 2

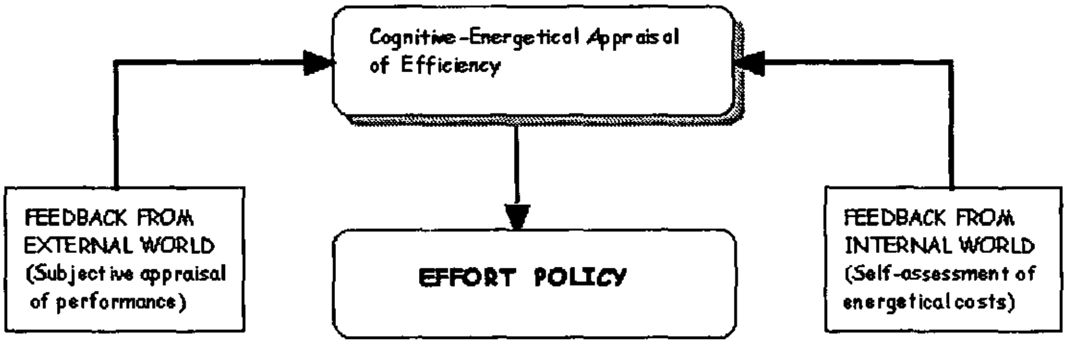


Figure 4. Schematic representation of the mental effort regulation model.

1.4 Structure of the thesis

This document constitutes a systematic extrapolation of the model of mental effort regulation described in the previous section. The purpose of the main literature review in Chapter 2 is twofold: (a) to describe theoretical perspectives on mental effort, and (b) to elaborate the model of mental effort regulation. With respect to the latter, those sources of external and internal feedback are described with reference to the literature. In addition, the assessment of cognitive-energetical efficiency and its influence on effort policy (i.e. the decision whether or not to invest mental effort) is also described in detail.

A number of empirical hypotheses concerning the model of effort regulation are stated in Chapter 3 which represents an introduction to the empirical work contained in the four following chapters. In addition, Chapter 3 is concerned with a brief review of methodological issues such as quantification of mental effort and testing environment.

Each empirical chapter describes an experiment to study different facets of the proposed model of effort regulation. The initial study (Chapter 4) represents an investigation into individual differences with respect to performance capability, a biological stressor (sleep deprivation) and mental effort regulation. The aim of this laboratory experiment is to study the hypotheses that: (a) good performance under strenuous conditions is the product of increased effort investment and (b) good performers are characterised by increased levels of effort investment and an awareness of external feedback.

The following studies represent applied research into the area of driving performance and driver impairment. An important area for the second study in Chapter 4 was the differential level of feedback awareness, which may exist between individuals. This theme is pursued in the fourth study (Chapter 5) where subjects are exposed to a range of biological stressors (e.g. full/partial sleep deprivation, alcoholic intoxication) in combination with sustained task activity. This study will address how different categories of stressors exert an influence on effort regulation and the appraisal of internal/external feedback.

It is assumed that effort may be poorly regulated due to the erroneous appraisal of external/internal feedback. This hypothesis represents the focus of the final study (Chapter 6) where subjects are provided with objective performance feedback during a simulated driving task. It is anticipated that the presence of objective feedback may directly influence effort regulation by highlighting external cues.

The final study (Chapter 7) will how effort policy and effort regulation is shaped by the joint influence of task demand and time-on-task. It is assumed that increased task demand will emphasise external factors and promote effort investment. On the other hand, increasing time-on-task will accelerate the rise of energetical costs and lead to effort reduction or conservation.

The results of all empirical data and their implications for the model of mental effort regulation are described in Chapter 8. A description of future research and the limitations of the thesis research are also included in this final chapter.

2 MENTAL EFFORT, PERCEPTUAL-MOTOR PERFORMANCE AND EFFORT REGULATION

2.1 Concepts of mental effort

The investment of effort represents a direct interface between energetical modification and ongoing cognition activity. This connection may be described in terms of a gain function (Wickens, 1986b) (Section 1.1). This modulatory link between mental effort and cognition is entirely descriptive, based on a simplistic metaphor where the rate or quality of cognitive activity is controlled by the level of effort investment, in the same way as the rate of cooking is controlled by the temperature of the oven. This metaphor may be highly intuitive but it is an inadequate heuristic to drive detailed analysis.

Mental effort is a difficult concept to define because it occupies a unique niche between biological and computational traditions in psychology. Effort is biological in the sense that it is influenced by energetical variables such as sleeplessness, drugs and stressors. On the other hand, if effort is capable of influencing cognitive activity, it must also be amenable to a computational conceptualisation. This gap was one of the subjects discussed in a paper by Hockey, Coles, & Gaillard (1986) written to highlight the discontinuity between 'wet' biological and 'dry' computational traditions in psychology and to unite both approaches within the hybrid discipline of energetics. These authors produced a wet-dry continuum of energetical constructs to illustrate a transition from the biological tradition to contemporary models of cognition. The original analysis of candidate mechanisms produced by Hockey, Coles, & Gaillard (1986) has been slightly adapted and appended in Table 1 below.

Table 1. Energetical concepts and their major influences in psychological theory. CNS = central nervous system, IPS = information processing system. Adapted from Hockey, Coles, & Gaillard, (1986).

Mechanism	Theorists	Principle Locus of Action
fuel	Thorndike	CNS
arousal	Malmo, Duffy	CNS
resources	Kahneman, Navon & Gopher, Wickens	CNS, IPS
Automatic/controlled processing	Shiffrin & Schneider	IPS

The following sections will describe each of the candidate mechanisms shown in Table 1 in more detail. In addition, each is considered as a candidate mechanism to represent the cognitive-energetical interface between performance and effort.

The first mechanism is described in terms of a fuel metaphor, where the brain functions as the engine for cognitive activity and mental effort is the fuel necessary for cognition. This is a simplistic analogy, which originated from early work on mental fatigue (Thorndike, 1900), but the metaphor is not without a certain logic. For example, it is known that the brain consumes 15% of cardiac output and accounts for between 20 and 30% of resting metabolic rate (Anderson, 1981). Despite these ample demands and in contrast to other organs, the brain has no local storage form for nutrients. Glucose cannot be stored to initiate energy release and therefore approximately 60% of body glucose from the liver goes directly to the brain (Van Toller, 1983). Therefore, the brain is totally dependent on a continuous supply of oxygen and glucose from the blood. This dependency is underlined by the fact that coma ensues if the brain is deprived of blood for as little as 10 seconds (Anderson, 1981).

At the cerebral site, it is estimated that 60% of energy is dedicated to synaptic activity (Madsen, 1993). A primary source of energy expenditure is the active pumping of ions in and out of the axon against the electrochemical gradient, following the transmission of an action potential (i.e. a nerve impulse). This process is necessary to re-establish the membrane potential and to optimise synaptic conditions for future transmission of action potentials (Madsen, 1993).

In basic terms, it is logical to assume the brain acts as both the engine and the site of action for cognitive activity. A fuel mechanism implies that increased procedural effort may be manifested as accelerated cerebral metabolism, increasing synaptic efficiency and thereby increasing the rate or quality of cognitive operations. In fact, several theorists have postulated that the influence of fatigue on performance may be interpreted in terms of cerebral metabolism, i.e. the energy is not available to completely pump ions from the synapses, therefore sub-optimal signal-to-noise ratios are produced at the synaptic gap (Crawford, 1961; Tsaneva & Markov, 1971). There is some empirical evidence to support this hypothesis (Mayleben, et al., 1998). However, the assumptions that: (a) all cognitive activity has a physiological analogue in the brain, and (b) the intensity of cognitive activity has a linear relationship with the processes of cerebral metabolism, represent a considerable conceptual leap beyond these fundamental connections (Beatty, 1986).

The concept of arousal has been used to describe a continuum of central nervous system (CNS) activity, ranging from deep sleep to high excitement (Duffy, 1962; Malmo, 1959). It has been proposed that both high and low extremes of arousal have a detrimental influence on human performance and this relationship is encapsulated in the Yerkes-Dodson Law (Yerkes & Dodson, 1908). The relationship between arousal and mental effort is ambiguous. It has been assumed that the arousal and effort overlap in terms of their influence on performance. Therefore, either an increase of effort or arousal may be sufficient to increase the intensity of behaviour (Kahneman, 1973). On the other hand, Sanders (1983) proposed that effort functioned as a modulator of arousal. In a recent paper, Hockey (1997) hypothesised that increased sympathetic nervous activity (i.e. increased arousal) may be a side effect associated with mental effort investment.

An arousal mechanism is based on the notion that mental effort influences performance by galvanising sympathetic nervous activity. This position is logical and is connected with the fuel mechanism described earlier. For example, increased sympathetic activity includes the release of catecholamines (in the form of adrenaline and noradrenaline) from the adrenal cortex and the adrenal medulla respectively. Adrenaline has the net effect of mobilising glucose as an energy source and increasing sympathetic nervous activity (Cox, Cox, & Thirlaway, 1983; Wesnes & Warbuton, 1983) and therefore, influencing the efficacy of cerebral metabolism.

Despite this logic, the arousal mechanism is beset with various problems concerning theoretical development, operationalisation and interpretation. The analyses of (Mulder (1986) and Hockey (1997) both contain an explicit distinction between effort and arousal. For example, if we are extremely tired, effort may be invested to compensate for detrimental influence of sleepiness on performance. If we are highly stressed, effort may be invested to aid concentration and to overcome anxiety. Both constructs have a negative correlation at extremes of arousal and a positive relationship within the medium range. On this basis, arousal is rejected as a potential mechanism for mental effort.

The topic of arousal and its relationship to mental effort is too complex to warrant a full treatment in the current thesis. The interested reader is directed to the following literature for in-depth treatment of arousal theory (Hamilton, Hockey, & Rejman, 1975; Hancock, 1988; Kahneman, 1973; Teigen, 1994; Venables, 1984), measurement issues concerning the arousal concept (Broadbent, 1971; Hockey & Hamilton, 1983; Lacey, 1967; Thayer, 1989) and the relationship of arousal to mental effort and related psychological concepts (Davies, 1983; Kjellberg, 1977b; Parasuraman, 1983).

The resource metaphor has been developed over twenty years to explore fundamental aspects of selective and sustained attention. Resource theory is based on a principle of scarcity, i.e. that our capabilities to perceive, analyse and respond are finite, both with respect to instantaneous task demands and sustained demands over time. If instantaneous or sustained demands exceed available resources, performance will decline as a direct result. The resource metaphor originated from concrete analogies to describe attentional limitations, including communication channels (Broadbent, 1958) and computer RAM (Moray, 1967). Later theorists such as Kahneman (1973) and Norman & Bobrow (1975) emphasised the singular, amorphous nature of processing resources. This concept had certain parallels with mental effort as both resources and effort were assumed to function as an undifferentiated, central source of cognitive modulation. In addition, Navon & Gopher (1979) introduced an explicit economic metaphor to describe performance trade-offs and utility assessments that may result from inadequate resource availability.

The investment of resources was closely linked to the investment of mental effort within several analyses, e.g. Kahneman (1973). According to Norman & Bobrow (1975), the relationship between resources (i.e. effort) and task performance could be classified into two archetypal forms. In the first instance, an improvement in performance may require a proportionate investment of resources or effort. This relationship describes a situation where the upper limit on performance is determined by resource limits. This pattern may characterise performance on a novel task, where increased resource investment promotes learning, and learning improves performance. In an alternative scenario, the investment of resources or effort may initially improve performance then reach an asymptote. At this point, further effort investment has no influence on performance. This scenario may occur if an operator attempted to read a degraded visual display, i.e. the upper level on performance is placed by sensory or psychophysical limits that were impossible to overcome. In this situation, the limits on performance and relationship between resources/effort and performance were termed data-limits. This scheme is illustrated in Figure 5 below.

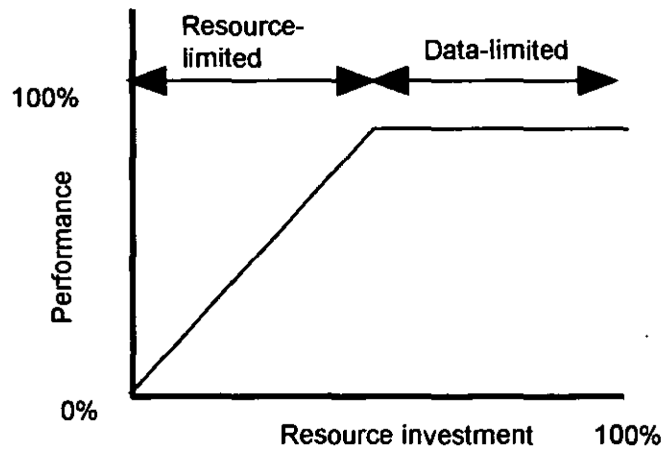


Figure 5. The relationship between performance and resource investment as described by Norman & Bobrow (1975).

The fundamental claim underlying the resource/data-limited distinction was that mental effort investment would not always improve performance. For example, if we are asked to add 2+2 on repeated occasions or read a newspaper under conditions of low illumination, increased mental effort investment will have no impact on performance.

Later theoretical developments were dominated by multiple resource frameworks (Wickens, 1980; Wickens, 1984), which emphasised a diversity of modality-specific resources. These theories represented a shift away from generalised, amorphous, processing resource and mental effort, which was historically associated with the single resource model (Kahneman, 1973). The debate between multiple and general resource concepts has occupied a substantive portion of the literature over the last fifteen years with no obvious advantage to either camp. Wickens (1991) reviewed the literature on resource theory and drew the following balanced conclusion: "To the extent that most variance in dual-task decrements is accounted for by structural factors and their interaction with demand factors, multiple resources become a more important construct. To the extent that most variance is accounted for by the main effects of task demand, a general capacity concept becomes more important" (p. 32). In other words, the type of resource theory that gains acceptance may be determined by the generality or specificity of resource scarcity.

This discussion of resource theory fails to shed any light on how mental effort may interface with cognitive operations. This may be due to the fact that hypotheses describing the modulatory influence of resource investment on performance are as imprecise as the gain function postulated by Wickens, (1986b) to describe effort.

A final fundamental problem with the resource framework concerns the operationalisation of processing resources. It is often assumed that resource investment is represented by performance variables such as dual-task interference. This presents a problem to the thesis as mental effort is based on a consideration of both covert and overt aspects of performance via the concept of efficiency (Schonpflug, 1985). The identification of resources with performance was criticised as being highly circular by several researchers (Navon, 1984). This controversy may have prompted the usage of psychophysiological (PP) measures as an alternative index of resource investment, e.g. heart rate variability (Mulder & Van Der Meulen, 1973), pupil diameter (Beatty, 1982b), evoked cortical potentials (ERPs) (Donchin, 1984).

This shifting operationalisation between cognitive conceptions and PP measurement landed the effort concept in the "no-man's land" between biological and cognitive traditions (Table 1). Ostensibly, resource models are proposed as cognitive theories with a particular emphasis on attention and dual-task performance. However, empirical work has employed PP measurement and the resource notion is entangled with arousal theory, whilst mental effort is entwined with both. Therefore, it is concluded that the resource concept may be identified with mental effort, but this association does not provide any additional insight or analysis into the mechanism of influence. The interested reader is referred to articles by Gopher & Donchin (1986) and Kramer & Spinks (1991) for in-depth treatment of the resource concept.

Cognitive investigation into visual search and memory has yielded a distinction between automatic and controlled information processing (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). These authors performed a programme of experiments using a multiple frame search procedure, where subjects searched for target items from an array of distracters. (Schneider & Shiffrin, 1977) and (Shiffrin & Schneider, 1977) varied several aspects of paradigm in a series of experiments. The important manipulation for the purpose of the current discussion was the relationship between the memory-set items and the items which appeared as distracters. This relationship was systematically varied such that there was a consistent mapping (CM) between memory-set and distracters, i.e. no items which appeared as targets ever appeared as distracters, or a variable mapping (VM), i.e. distracter items occasionally appeared as targets and vice versa. This manipulation had a profound effect on the accuracy of performance. The CM condition produced high accuracy scores whereas performance on the VM task was much poorer. The former was identified with automatic processing whereas the latter was termed controlled processing.

Several aspects of the automatic/controlled processing dichotomy are significant for procedural aspects of mental effort. Automatic processing is described as fast, error-free, capable of functioning in parallel and relatively effortless. On the other hand, controlled processing is slow, error-prone, capable only of serial processing and is relatively effortful (Fisk, Ackerman, & Schneider, 1987). The implications for mental effort are twofold: (a) effortful processing is evoked when the features of target stimuli must be explicitly checked under VM conditions, and (b) the process of memory retrieval may raise the mental effort associated with the VM condition as targets must be updated, encoded and retrieved on a regular basis.

This section has travelled along a wet-dry continuum from biological concepts to cognitive mechanisms (Table 1). On the basis of this consideration, it may be concluded that the biological traditions tends towards imprecise definition and operationalisation, whereas the cognitive perspective fails to explain the influence of energetical variables such as fatigue and stress on performance (Hockey, Coles, & Gaillard, 1986). In addition, the wet-dry continuum begs a question as to whether we are considering alternative mechanisms for effort modulations or merely different levels of the same mechanism?

The fuel metaphor was based on cerebral metabolism and it is logical that cerebral metabolism may constitute a biological substrate for controlled processing[1]. If an individual is forced to increase his or her levels of controlled processing, it may be assumed that fuel consumption will increase at the synapses. If a higher level of effort is sustained, sympathetic activity may increase in the CNS, which may boost psychological arousal. The purpose of this activity may be increase fuel supply to the brain via the hormonal mechanisms described earlier. However, the increased rate of fuel expenditure combined with sympathetic dominance may be associated with certain costs (Hockey, 1993; Hockey, 1997; Schonpflug, 1983). These costs may include heightened tension due to increased sympathetic dominance.

This section has described how various concepts related to mental effort may be reconciled in order to influence performance. It was argued that effort may manifest itself as an increase of controlled processing, buttressing performance by increasing the frequency of attentional checking to task activity and therefore, aiding concentration on task performance. However, increased controlled processing demands higher levels of cerebral metabolism, which in turn accelerate energetical costs such as stress and fatigue across time-on-task. This link between controlled processing and energetical variables represents a manifestation of the antagonistic dynamic between the need to perform and the desire to withdraw (Section 1.2). This dynamic underlies the requirement for mental effort regulation discussed in the following section.

2.2 Mental Effort Regulation and the Formulation of Effort Policy

This section is concerned with a focused literature review to describe a model of mental effort regulation. This review will begin by describing and reviewing the theoretical lineage represented by homunculus theories of hierarchical control. According to the proposed model, the regulation of mental effort is driven by task-related and energetical feedback.

---

[1] Although it is only fair to add that cerebral metabolism may constitute a biological substrate for any type of cognitive activity.

2.2.1 The Homunculus

The conceptualisation of mental effort as a ubiquitous central capacity (Kahneman, 1973) brought the concept into alignment with hierarchical models of cognitive control. These models may be termed "homunculus" theories that posit the existence of an executive controller exerting widespread influence over psychological activity. This controller has been caricatured by critics as the GPLCCP (General Purpose Limited Capacity Central Processor) (Allport, 1980) or the little man in the brain (Brown, 1988). The homunculus models of executive control represent alternative mechanisms for mental effort regulation as both are associated with the exercise of ubiquitous control.

The prototype for the homunculus was the Test-Operate-Test-Exit (TOTE) concept inspired by cybernetics and devised by (Miller, Galanter, & Pribram, 1960). These authors postulated that behavioural control was regulated via iteration between Operate activities and Test procedures. An important feature of the TOTE unit was its inherent flexibility. TOTE units could operate in a linear chain, i.e. the exit stage of TOTE *a* may act as a cue to produce the initial Test on TOTE *b*, alternatively, the operate phase could expand hierarchically to incorporate subordinate TOTEs. The hierarchical linkage between various TOTE units produced a distinction between superordinate 'parent' TOTE units (i.e. controllers) and subordinate 'child' TOTEs (i.e. operators).

The homunculus as an upper level controller made its formal debut in the area of experimental psychology. Broadbent (1963) discovered that arousing and de-arousing stressors did not interact in an additive fashion with respect to performance, i.e. barbiturates failed to ameliorate the influence of sleeplessness. This discrepancy was the central topic of his book "Decision And Stress" (Broadbent, 1971) which introduced an initial formulation of a two-tier control structure. Within this model, a lower mechanism sustained well-established decision processes with the intermittent assistance of an upper mechanism. The purpose of the latter was to monitor the performance of the lower mechanism, and to buttress performance under both extremes of the arousal continuum. Therefore, a non-additive interaction between stressors could be explained as different stressors have a specific impact on either the upper or the lower mechanism. The two-tier model proposed by Broadbent (1971) is illustrated in Figure 6 below. This concept was refined some years later (Broadbent, 1977) without any explicit link to stress and performance. However, the main characteristic of regulation of routine activities by a superordinate controller was sustained.

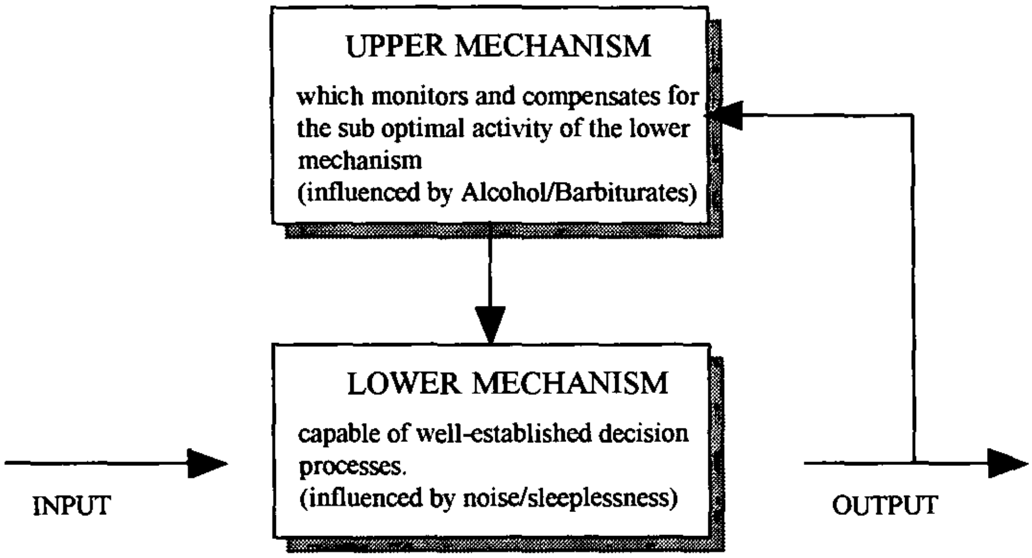


Figure 6. Two-tier model proposed by Broadbent (1971).

The executive controller hypothesised by Broadbent was also included in the model of working memory devised by Baddeley & Hitch (1974). This model contained three components, two slave systems (Phonological Loop and Visuo-Spatial Sketch), each of which were co-ordinated by a Central Executive. The purpose of the latter was to select, co-ordinate and plan the activities of the slave-systems, but these executive functions were not described at any level of detail in the early versions of the model.

The essential aspect of both models is the identification of mental effort with the activity of the upper mechanism or central executive. Behaviour at the lower level was characterised as relatively effortless whereas effortful control was the domain of the upper level controller.

These ideas were supplemented by the notion of a unitary, "all-purpose" executive controller described by Kahneman (1970; 1973). In addition, Kahneman (1973) was the first to bring the notion of effort to prominence as the source of executive control. Kahneman's model included three important theoretical innovations: (a) the central executive was characterised as a unitary and finite source of 'fuel' that was synonymous with mental effort, (b) the amount of available effort was influenced by psychophysiological extremes of sleepiness and stress, and (c) distribution of this resource throughout the cognitive system was determined by an allocation policy. This conception produced a fluid homunculus, which could be simultaneously characterised as a unitary energy source and a centralised point of energy distribution. In addition, Kahneman made an explicit assumption that the executive exerted an influence over lower level activities via the regulation of mental effort.

This theme of higher-level control was continued in other theoretical domains such as action selection and control. For instance, in his description of the dominant action system, (Shallice, 1978) argued that action selection may be modulated by a higher-level system known as the Supervisory Attentional System (SAS). The function of the SAS was to provide an executive input into action planning and action selection (Norman & Shallice, 1986; Shallice, 1982; Shallice, 1988). The link between mental effort and the activity of the SAS is inferred on the basis that those factors responsible for SAS intervention are also associated with increased effort, e.g. planning/decision-making, error-correction, novel sequences of actions, a degree of danger, and the overcoming of a strong habitual response (Shallice & Burgess, 1993).

The hierarchical model of control developed by Hockey (1986a;1993; 1997) represented an extension of the previous models described by Broadbent and Kahneman. The latest version of Hockey's model (Hockey, 1997) is shown in Figure 7. This model proposed that task goals are relayed to the action monitor via the supervisory controller. The activity of the lower loop concerns routine/effort-free performance as dictated by the goals relayed from the supervisory controller. If these goals cannot be maintained by the lower loop (Loop A), performance discrepancies are detected via the Effort monitor and relayed back to the Supervisory controller. The modulation of task performance via the upper control loop (Loop B) involves one of two strategies, either effort investment (to increase task performance levels) or effort conservation (to strategically reduce task goals to be sustainable with minimum mental effort).

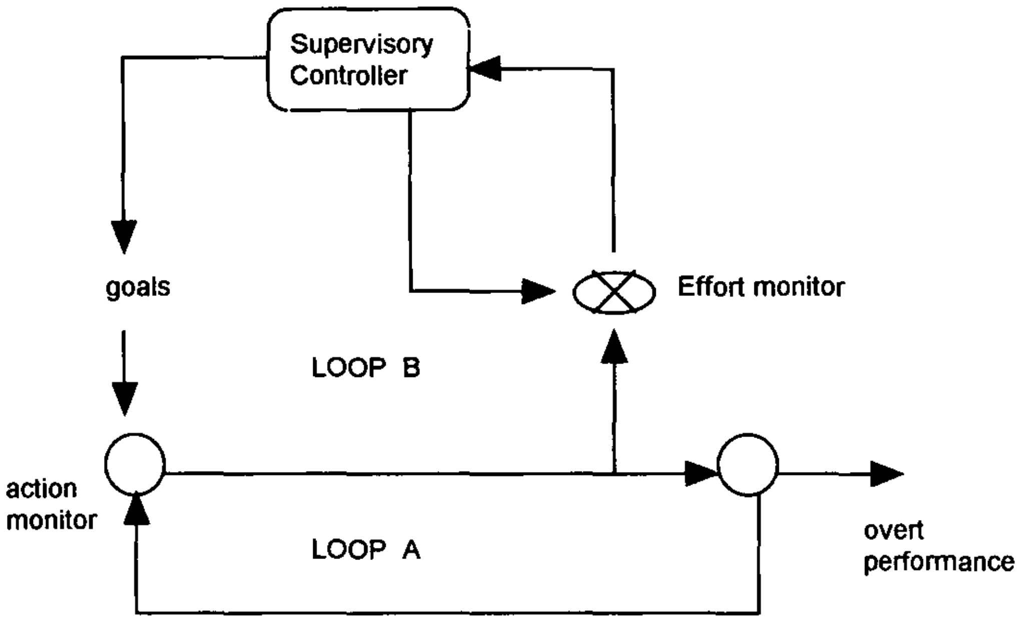


Figure 7. The effort regulation model devised by Hockey (1997). Loop A represents routine activity whereas Loop B shows effortful, adaptive control.

The model proposed by Hockey (1997) illustrates an interaction between top-down supervision and bottom-up discrepancy reduction. Hence, the regulation of mental effort results from the iteration between the Loop B and Loop A in Figure 7 – supplementing upper-level, top-down control with autonomous control subroutines from the lower level. A similar framework was described by Lord & Levy (1994) who also included diverse modes of control within their model of self-regulation. Specifically, Lord & Levy (1994) contrasted top-down control due to goal activation with bottom-up control via discrepancy reduction., i.e. control originating downwards from goals or travelling upwards due to the occurrence of errors. These authors postulated that both modes of control were necessary to "help maintain one's current cognitive frame as well as a specific, implemental mind-set" (p. 350). This iterative model of top-down and bottom-up control places equal emphasis on each and may be termed a behavioural heterarchy.

Nevertheless, all homunculus theories largely subscribe to a hierarchical perspective of behavioural regulation. Within this framework, the regulation of mental effort is a top-down activity. Therefore, the upper level or executive is the source of goal development and responsible for an appraisal of costs and performance quality. The purpose of the hierarchical design is to delegate control in order to limit the involvement of the upper mechanism/superordinate controller. This strategy of control is justified on economic grounds, i.e. it is assumed that upper-level control via Loop B is inherently more effortful than control within Loop A.

Despite a long theoretical lineage, the homunculus remains a controversial theoretical entity. According to Allport (1980), postulation of an executive controller falls into a trap of infinite regress as it begs a question with respect to the ultimate level of control (i.e. what controls the controller?). In addition, it is difficult to operationalise these theories in a fashion that clearly differentiates upper and lower levels of control. For example, if an upper mechanism expresses itself via the modulation of lower levels of activity, how may one level of behavioural control be distinguished from the other?

Allport (1980;1992) has argued that the heterogeneity of attentional mechanisms constitutes an argument against the existence of a GPLCCP. Furthermore he claimed that the concept of a GPLCCP was too imprecise and therefore "provides an inadequate heuristic for driving research" (Allport, 1980) (p. 143). An alternative and more moderate approach to the same problem was outlined by Baddeley, (1996). This author accepted the existence of the homunculus as a convenient piece of fiction, but recommended its retention, both as an organising principle and to represent the "problem" of central control. This strategy provides the homunculus with a stay of execution, but only to be dissected by further research until it may be declared obsolete.

2.2.2 The purpose and limitations of mental effort regulation

This section will revise and expand the model of mental effort regulation illustrated in Figure 4. It has been argued that mental effort is finite and therefore, must be allocated according to rationalist or economic principles. The purpose of a mechanism for mental effort regulation is:

- To monitor cumulative levels of mental effort expenditure and reserves in order to predict surplus and deficit
- To monitor feedback from external task sources to estimate requisite level of effort to sustain acceptable levels of task performance
- To monitor feedback from internal energetical sources to estimate costs necessitated by continued effort expenditure
- To formulate effort policy, whether to invest or conserve effort correspondent with two criteria: (a) avoid exhaustion of effort reserves, (b) maintain adequate performance, and (c) avoid unacceptable levels of task-related stress, fatigue or discomfort.

The formulation of effort policy[2] is determined via dual-feedback. However, feedback may only be interpreted with the aid of 'standards' or criteria, which may be derived on the basis of experience. For example, an experienced operator may anticipate a performance quality of x and a psychophysiological status of y. This interaction represents a crucial transaction between standards that are subjectively derived and events in the external world (Frese & Zapf, 1994).

The introduction of subjective self-appraisal into this process brings a degree of potential bias and distortion. The individual can register feedback from either external or internal sources, but his or her appraisal of feedback may be neither completely representative nor entirely accurate. For instance, there is a wealth of evidence from the study of human error that the act of subjective appraisal is associated with inherent limitations, e.g. selectivity, overload, perceptual confusion, incorrect inferences etc. (Reason, 1990).

This transactional aspect of appraisal was encompassed by the framework for the study of stress described by Lazarus & Folkman (1984). According to these authors, an individual may appraise external events in terms of their personal significance and capacity to react to them. On this basis, Lazarus & Folkman (1984) proposed that appraisal took a primary and a secondary form. The former corresponded to an evaluation of personal meaning and significance for individual wellbeing. Therefore, the direction of primary appraisal may be prioritised based on personal relevance and/or potential threat to wellbeing.

The secondary process of appraisal proposed by Lazarus & Folkman (1984) was concerned with available coping options to deal with potential sources of stress. This primary appraisal represents an assessment of external and internal feedback, whereas secondary appraisal concerns the formulation of a policy whether to invest or to conserve mental effort (Hockey, 1993; Hockey, 1997).

The key to appraisal within the current model of regulation is its essential subjectivity, i.e. an emphasis on personal significance, individual bias, attentional selectivity. This transactional perspective runs counter to the Objectivist model of volition proposed by Binswanger (1991) in Section 1.1, where effort regulation was based on a conceptual faculty operating in a glorious absence of selectivity,

[2] This term is chosen to represent an affiliation with the concept of an effort allocation policy as described by Kahneman (1973).

rrationality or emotional bias. Within the current framework, the formulation effort policy is letermined by a subjective appraisal of external/internal feedback, regardless of accuracy or :epresentativeness.

2.2.3 Feedback from external and internal sources

The background to this discussion is provided by two theoretical treatments of feedback control; homeostasis and cybernetics. The former was described by Cannon (1932) who outlined a series of physiological mechanisms dedicated to the maintenance of a stable cellular environment. This generic process of feedback and adaptation was expanded by Wiener (1948) into the multidisciplinary field of cybernetics. In his first book on the topic, Wiener (1948) constructed a persuasive thesis positing a generic process of feedback and regulation that was apparent in mathematics, communication theory, engineering control and psychology, as well as biological activity.

The negative feedback loop was the core element at the heart of both theories. An important distinction between the feedback loop concept as employed by Cannon and Wiener concerns the extent to which the perception of feedback may be regarded as either involuntary or voluntary. Cannon employed feedback to describe passive, involuntary regulation within the human nervous system, i.e. the regulation of the autonomic nervous system do not involve conscious control. On the other hand, Wiener applied the control loop metaphor in more catholic terms to emphasise its universality across disciplines, as a deterministic element within a mechanic device and a stochastic process within human decision-making.

It is hypothesised that feedback for effort policy may also originate from external and internal sources. For example, we may decide whether to continue with a particular task based on our perceptions of task quality or success, which originates in feedback from the external world. In addition, we may wish to consider internal feedback of our level of wellbeing, e.g. do we feel tired, sick or bored.

The significance of feedback for the mental effort regulation should not be underestimated. It is postulated that the perception of degraded performance or increased discomfort serve an essential requirement, i.e. providing cues with respect to the magnitude of desirable effort investment. In addition, if subjective discomfort is accelerated, these negative signals function as feedback that finite effort reserves are depleted or approaching exhaustion. Therefore, feedback from performance and the self serves a double-function within the proposed model: (a) to supply cues regarding the effectiveness of effort policy, and (b) to indicate remaining effort reserves. It is postulated that both functions of feedback are combined to constitute an assessment of utility, which is used to determine future effort policy. The sources of both external and internal feedback are described and reviewed in the following sub-sections.

2.2.4 Feedback from the external world

Feedback from the perceptual (external) world is the basis upon which we assess performance and the quality of performance. On some occasions, feedback may be very explicit, e.g. proliferation of overt errors or catastrophic episodes of task failure. In other cases, feedback from the external world may be cross-referenced with an internal performance standard for the purposes of assessment (Frese & Zapf, 1994), i.e. a comparison with a cognitive standard as hypothesised by Carver & Scheier (1981). The presence of internal, performance standards are crucial to the appraisal of feedback. These standards represent a range of expectations developed on the basis of past experience (Bandura, 1997). These standards are used to provide important contextual cues, to "benchmark" current performance and to classify output as normative, sub-normative or supra-normative.

Those internal standards representative of 'normative' performance may be complex and multidimensional. For example, current performance may be referenced against an array of standards associated with: (a) the frequency of errors, (b) rate of task progress, and (c) perception of physical strain associated with response production. In addition, performance feedback may serve as a secondary source of feedback for the perception of task demand, i.e. increased error, sub-standard performance, slow progress or physical strain may be appraised as indicators of increased task demands. This latter aspect was investigated by Hancock (1989) who studied whether task-naïve subjects could distinguish variations in task load based on repeated experiences of performance failure. An analysis of subjective workload measures indicated that subjects perceived increased workload (particularly due to frustration) when performance was unsuccessful.

The occurrence of error represents a common mode of performance discrepancy and feedback to the individual. The presence of an error may be detected by the individual in one of three ways: (a) discovered via self-monitoring, (b) an overt signal from the environment, and (c) feedback from a second individual (Reason, 1990).

Several studies of error correction during sequential activity were performed by Rabbitt (1981) to describe some limitations of self-monitoring. For example, in an earlier study, Rabbitt, Cumming, & Vyas (1978) made a distinction between false identification errors (i.e. indicating that a target was present when it was not) and errors of omission (i.e. failing to respond to a target that was present). They found that omissions were more common than the false identifications. In addition, more omission errors were detected than errors associated with false identification. These findings illustrated that individuals may recover from errors associated with inappropriate response selection, whereas perceptual errors were more difficult to detect/correct.

This work was supplemented by a study of skilled typists, who were asked to cease typing as soon as an error was detected Rabbitt (1978). The results of this study revealed that typists were capable of detection within one or two keystrokes of error commission (approximately 182 msec - see Logan, (1985) for discussion and extensions to this area of research). Therefore, error detection may be associated with differing levels of demand, but once detected, individuals are capable of a rapid response.

The limitations of self-monitoring are apparent within the GEMS (Reason, 1987) where action slips may occur (i.e. inappropriate actions-not-as-planned) if attentional checks are not performed at the appropriate times. This aspect of error production has been analysed by Norman (1981) and Baars, (1988) with respect to action control. According to the latter, action slips are indicative of an intermittent mode of self-monitoring, where routine aspects of action control and performance are handled unconsciously and "non-routine choice-points" (Baars, 1993) (p. 281) are subject to attentional checking. Therefore, the individual may only monitor performance for discrete periods within a stream of response output.

The process of performance monitoring is complicated during problem-solving activities at the strategic level such as plan generation/selection. Allwood (1984) studied error detection during statistical problem solving using verbal protocol analyses. This investigation suggested three prototypical episodes of error detection: (a) standard checks initiated as a general check on progress, (b) error suspicion when the subject was aware of a discrepancy or suspected the presence of a discrepancy without identifying the source of the error, and (c) direct error hypotheses formation when the subject made an abrupt detection of a presumed discrepancy. These episodes of discrepancy detection were triggered by a range of variables, including a departure from subjective expectations. The main difference between the three detection episodes being that standard checks are self-initiated progress checks, whereas (b) and (c) are provoked via feedback from environmental cues. It has been suggested by Brichcin (1982) that a strategy of self-initiated checking during performance assessment may produce an active (i.e. effortful) mode of performance regulation, which may increase the fidelity of goal-feedback comparison.

A complimentary relationship between internal standards and overt feedback has been demonstrated throughout studies of error detection. In addition, self-monitoring of performance quality is an important source of information for the assessment of task difficulty or subjective mental workload. In this case, individuals employ performance feedback as a barometer of task demand and as a means of anticipating changes in task demand. Moray (1982) suggested that subjective perception of workload/task demands could be related to the rising probability of failure in the near future. This statement found some support from the study conducted by Hancock (1989) described earlier.

Other researchers have developed more quantitative conceptualisations of subjective workload. For example, Tulga & Sheridan (1980) presented a model where subjective workload was estimated based on task number, task pacing, task deadlines and human productivity. The temporal character of this model is echoed in recent research conducted by Hendy, Liao, & Milgram (1997) and the validated model proposed by Hancock & Caird (1993), where subjective workload is represented by two dimensions: perceived time for effective action and perceived distance from desirable goal. In all cases, the perceptions of the individual and the experience of subjective workload are translated into potential consequences for task performance (and therefore, the formulation of future effort policy).

The importance of the subjective workload concept as a source of regulatory feedback has been somewhat undermined by the inconsistent relationship between subjective and performance-based measures of mental workload (Derrick, 1988; Gopher & Braune, 1984; Yeh & Wickens, 1988). On the other hand, there is a good deal of evidence for correspondence between subjective workload and performance (O'Donnell & Eggemeier, 1986). If this were not the case, an appraisal of workload/task demand based on performance feedback would be utterly useless for the formulation of effort policy. Therefore, it is important for an effort theory to account for divergence between subjective estimates and actual performance (Muckler & Seven, 1992).

An analysis of subjective-performance dissociation was provided recently by Hancock 1996) based on a series of studies on tracking performance. Hancock (1996) presented his analysis in the form of a table, which is reproduced below (Table 2). The solid arrow moving left-to-right describes those instances of complete association, i.e. when high subjective workload is associated with worse performance and vice versa. The dashed arrow provides examples of complete divergence. Hancock (1996) pointed out that higher workload coupled to better performance may be indicative of successful compensation for increased task demand. On the other hand, the obverse cell when worse performance is associated with a reduction of subjective workload may indicate that the individual has effectively "given up" or adopted a lower standard of performance as a means of ameliorating increased workload (Hockey, 1997).

Cell (a) describes a dissociation scenario where stable performance is associated with increased subject workload. This divergence may occur when an individual is actively investing effort merely to achieve performance stability. Both cells (b) and (c) provide examples of dissociation when subjective workload does not respond to an increase or decrease of performance effectiveness. In both cases, the subjective appraisal of feedback by the individual is insensitive to changing performance quality. This may be due to opaque feedback cues (e.g. perceptual errors described by Rabbitt, Cumming, & Vyas, (1978)). The cell (d) illustrates how workload may decline whilst performance is sustained at a stable level. This pattern of dissociation may occur as the individual develops increasing skill based on task experience, i.e. may operate at a higher level of task efficiency. The matrix devised by Hancock (1996) provides an orderly, transactional analysis of the relationship between subjective appraisal and objective performance.

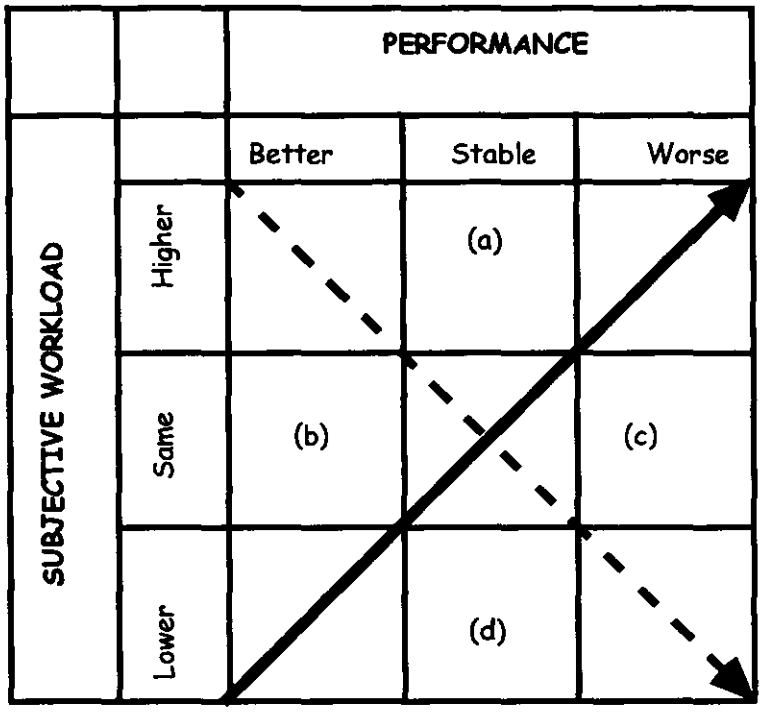


Table 2. A matrix to describe the relationship between task performance and subjective mental workload (Hancock, 1996).

It is hypothesised that the internal model or representation (Stassen, 1989) provides the context for a subjective appraisal of task performance. This internal model encapsulates task dynamics, task strategies, performance standards and associated expectations. A divergence between the internal model and actual task dynamics/demand may be responsible for those dissociations between subjective appraisal and performance described in Table 2. According to Hacker, Plath, Richter, & Zimmer, (1978), differences with respect to the quality of internal models (i.e. fidelity, representativeness) are important source of performance variability. These authors stressed the importance of internal models as a means to anticipate changing task demand and to formulate effort policy. The internal model may represent the only means of providing context for self-monitoring activities and the subjective appraisal of task feedback. The internal model may be the basis for performance standard definition, error detection and the appraisal of subjective mental workload. If the internal model is incomplete or poorly defined, it is difficult to produce an accurate appraisal of task feedback and to formulate appropriate effort policy.

The internal model may be the basis for standard formulation at all levels of the behavioural heterarchy and therefore, represents the foundation of all instances of discrepancy detection. This point is reinforced by Figure 8 below. This diagram is based on the closed-loop, causal model of workload developed via fuzzy measurement by Moray, King, Turksen, & Waterton (1987). These researchers produced a model of mental workload in order to contrast crisp and fuzzy workload measures. A fuzzy approach to workload measurement permitted formulation of a closed-loop model with multiple causal links between objective variables and subjective entities uncovered via regression analyses. The model proposed by Moray, King, Turksen, & Waterton (1987) is reproduced in Figure 8 with the addition of an internal model or task representation by the author[3]. This model occupies a juncture between objective task variables, the appraisal of task difficulty and consequent formulation of effort policy.

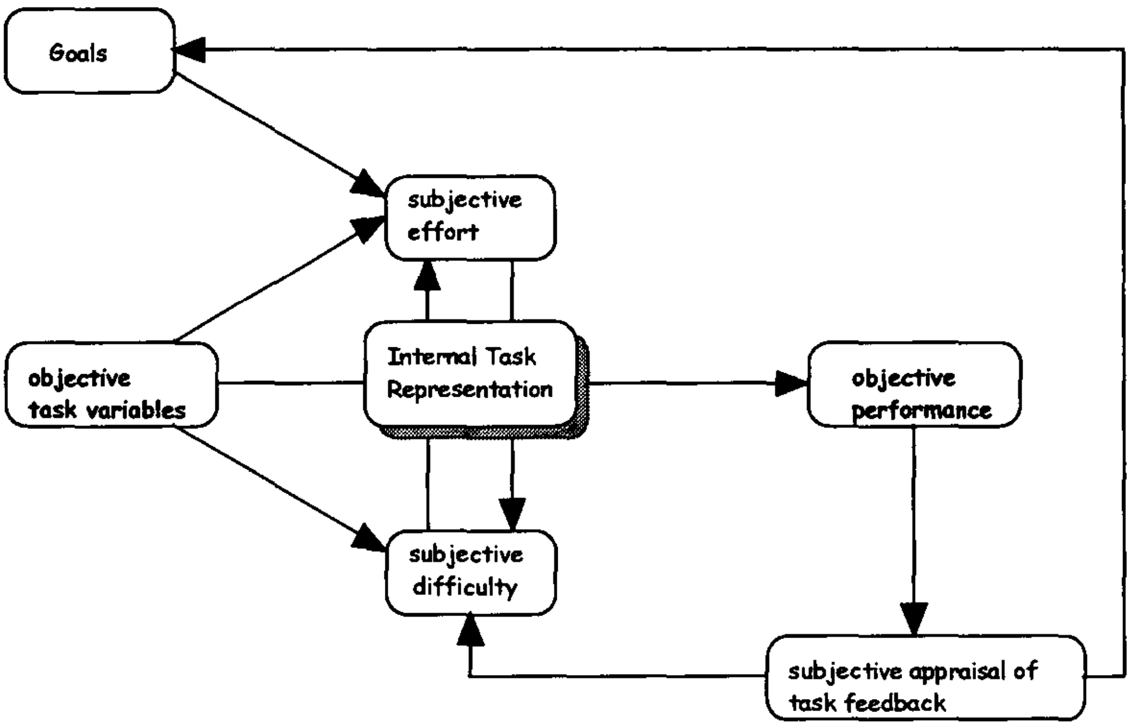


Figure 8. A modified version of the hypothetical, closed-loop causal model of workload proposed by Moray, King, Turksen, & Waterton (1987).

The model of workload shown in Figure 8 contains all those elements that constitute the performance feedback loop. Feedback from the external world is appraised in terms of task goals, which in turn, inform the level of subjective effort deemed necessary to perform at an appropriate level. The internal

[3] The original model produced by Moray et al (1987) is identical to Figure 8 in all other respects.

task representation provides a locus for the interaction between objective task variables and objective performance. In other words, it is formulated and re-formulated by an understanding of the relationship between task variables and their consequences for performance. In addition, the internal model encapsulates the interaction between subjective effort and estimates of task difficulty. The model of the task represents the juxtaposition between internal and external performance standards. It is hypothesised that the internal task representation provides the framework for an assessment of external, performance-based feedback.

2.2.5 Feedback from the internal world

When an individual performs a given task, it is hypothesised that performance is achieved with a number of cognitive-energetical costs (Hockey, 1993; Hockey, 1997). For example, the performance of a difficult task may be associated with increased sympathetic nervous activity leading to heightened alertness and/or the experience of stress. At the other extreme of an arousal continuum, task activity under conditions of extreme fatigue or sleep deprivation may be correlated with decreased energy and a number of characteristic physiological symptoms, e.g. heavy eyelids, yawning, muscular aches. These internal, cognitive-energetical factors are important because they may impinge on performance quality if effort is not invested to counteract their influence.

In Section 2.2.3, a distinction was made between reflexive and perceptual feedback loops with respect to physiological homeostasis and other manifestations of cybernetic control. It is natural to associate internal feedback with the reflexive process of homeostasis, however, this link may be overly simplistic. It is postulated that internal feedback has a direct association with homeostatic control when thermoregulation, blood pressure regulation and other essential, physiological control processes breakdown, e.g. when the limits of physiological adaptability are reached (Hancock & Warm, 1989). The effort regulation model employs the concept of internal feedback in much broader terms, to encompass a range of energetical costs originating internally, which may have an indirect association with task performance. These internal symptoms may be registered via reflexive or perceptual feedback loops. For example, overt symptoms such as yawning or high perspiration are noted automatically, whereas rising levels of irritation or reduced energy represent psychological symptoms that are perceived by the individual.

These energetical costs are related to the well-being of the individual, which in turn, define the level of discomfort accompanying task activity. In conceptual terms, the regulation of energetical costs may be associated with the comfort perceptual system described by Bartley (1970). This system was hypothesised to liase between the senses and the processes of physiological homeostasis in order to maintain psychological wellbeing, i.e. by monitoring and providing feedback with respect to awareness of experiential bodily comfort.

The comfort system corresponds to an internal feedback loop, responsible for monitoring and regulating those energetical costs associated with task performance. It may be hypothesised that cybernetic control is accomplished via the existence of standards or set-points with respect to internal feedback. In other words, any given level or duration of task performance is associated with an anticipated energetical cost, which acts as a standard for internal feedback.

A prominent source of internal feedback is physiological symptoms of discomfort, e.g. aches and pains. In a large survey of over one thousand students, Pennebaker & Skelton (1978) found that almost eighty percent reported at least one symptom from a list of twelve items, e.g. headache, upset stomach, sore muscles, chest pains, nasal congestion, watering eyes, ringing in ears, racing heart, dizziness, sweaty hands, flushed face, shortness of breath. It would appear that symptoms of discomfort are relatively common even among a healthy population. The surprising proliferation of physiological symptoms reported by Pennebaker & Skelton (1978) begs a question – to what extent are people capable of providing an accurate report of their internal, somatic states? Preliminary data reported by Pennebaker & Skelton (1978) revealed low to moderate levels of correlation, e.g. $\underline{r} = 0.28$ between the perception of nasal congestion and exhalation rate through the nose. This general finding was originally reported by Mandler, Mandler, & Uviller (1958) who distinguished individual traits with respect to over- and under-estimation of autonomic reactivity[4]. These data illustrated that the comfort system is essentially perceptual rather than reflexive, and therefore prone to those sources of bias' which characterise perceptual distortion in the external world, e.g. selectivity, psychophysical distortion.

The experience of fatigue under operational conditions has been studied intensively using subjective ratings, based partly on the frequency of characteristic, physiological symptoms. For example, the

[4] This trait of differential, autonomic sensitivity was later conceptualised as private body consciousness by Miller et al (1981).

symptom checklist devised by Yoshitake (1971;1978) distinguished: (a) general drowsiness/dullness, (b) difficulty concentrating and (c) specific physiological symptoms as components of operational fatigue. These factors may constitute energetical/compensatory costs (Hockey, 1993; Hockey, 1997) during sustained task performance.

The use of symptom checklists was extended to the study of driver fatigue during prolonged journeys by Nelson (1987), Desmond, Matthews, & Hancock (1997) and Nilsson, Nelson, & Carlson (1997). In a study of sustained, simulated driving, Nelson (1987) asked subjects to drive until they felt unable to continue. He reported that fatigue symptoms such as backache, headache and tired eyes increased in a linear fashion with time-on-task. However, certain severe symptoms such as upset stomach and dizziness were observed to increase just prior to the subject decided to discontinue the driving task. Therefore, qualitative groups of symptoms may be assessed in terms of differential severity. This approach was developed by Desmond, Matthews, & Hancock (1997) who performed a factor analysis in order to distinguish four groups of fatigue symptoms: (a) visual fatigue, (b) malaise (e.g. illness), (c) boredom, and (d) muscular fatigue. This factor analysis emphasised the role of psychological symptoms (e.g. boredom) alongside traditional physiological symptoms.

Affective changes in mood represent another source of cognitive-energetical costs associated with task performance. Task activity may have a profound effect on mood in terms of subjective alertness, tension and positive/negative affect. Traditional approaches to mood measurement have involved the construction of a self-assessment tool known as the mood adjective checklist , e.g. (Nowlis, 1965). This tool was developed by Thayer (1967) to support a two-dimensional conceptualisation of psychological arousal, which distinguished two bipolar components of arousal, one concerning alertness and another concerning tension. The development of this work, both in terms of theoretical progress and empirical studies, is described in detail by Thayer (1989).

The development of the mood adjective checklist was advanced by Matthews, Jones, & Chamberlain (1990) who developed a subjective tool which subsumed three constructs: energetical arousal (alert-tired), tense arousal (relaxed-tense) and hedonic tone (happy-sad). The relationship between mood and performance was studied extensively by Matthews and his colleagues. In a recent summary of their work, Matthews, et al. (1997) contrasted the influence of a vigilance task, e.g. sustained attention over 48 minutes (N=229), and a working memory task, e.g. memorisation of digit strings combined with counting for 12 minutes (N = 151) on the three components of mood. It was found that vigilance significantly reduced energetic arousal and to slightly reduce hedonic tone, but had little effect on tense arousal. Performance of the working memory task was associated with increased energetic and tense arousal, but exerted little influence on hedonic tone. These findings are indicative of increased sympathetic nervous activity associated with effortful working memory activity, creating an increase of alertness and tension, e.g. compensatory costs (Hockey, 1993; Hockey, 1997). A detailed consideration of the relationship between mood and performance may be found in Matthews (1992).

A range of energetical costs associated with task performance was operationalised in the form of a comprehensive stress state questionnaire devised by Matthews, et al. (1997). This questionnaire is composed of ten constructs, which have been factor-analysed into three meta-constructs: engagement, distress and worry.

The first meta-construct of engagement provides an index of "the capacity of the task to provoke a *commitment to effort* and application" (Matthews, et al., 1997) (p. 13). This construct of engagement is composed of three scales, energetic arousal (as described in the previous paragraph), task motivation and the level of concentration. (Matthews, et al., 1997) correlated engagement scores with several indices of performance collected from different experimental studies. Data from two exemplar studies are included in the current discussion: (a) the working memory task described in the previous paragraph, and (b) a successive vigilance task, e.g. sustained attention combined with a load on working memory (Warm & Dember, 1998). It was found that energetic arousal had a positive relationship with task performance for both tasks. In addition, the level of concentration exerted by the individual was a significant predictor of vigilance performance. Therefore, reduced engagement constitutes a distinct source of energetic cost, resulting in reduced energy associated with declining motivation and concentration.

The second meta-construct described by Matthews, et al. (1997) was called distress which relates to "*overload of capacity* to perform successfully" (p. 14). This factor is related to the transactional definition of stress defined by Lazarus & Folkman (1984). An individual suffers from symptoms of distress when he or she feels that adequate performance may be beyond current capability. This meta-construct is composed of three constituents: tense arousal, hedonic tone (both were described earlier in this section) and confidence. Matthews, et al. (1997) reported that hedonic tone showed a significant correlation with task performance during both working memory and successive vigilance activities. Therefore, individuals who perform to an adequate level experienced positive affect and vice versa. In addition, it was found that confidence was a significant predictor of working memory performance, but this relationship was not significant during successive vigilance. The experience of distress may be associated with performance under demanding conditions. Based on the analysis provided by Matthews, et al. (1997), energetical costs would manifest themselves in a distressed individual via increased tension, negative affect and declining confidence in the ability to perform.

The third meta-construct of worry is related to a theoretical model of self-directed attention devised by Wells & Matthews (1994), known as the Self-Regulatory Executive Function (S-REF) model. In brief, the S-REF model postulates a hierarchical cognitive architecture composed of three levels: (1) automatic/reflex processing units, (2) controlled/voluntary processing units, and (3) stored knowledge and self-beliefs. The second level of architecture is responsible for monitoring low-level activity at level (1) and to liase with high-level knowledge concerning self-beliefs (3). The model assumes that monitoring activity at level (2) encompasses a range of potential discrepancies, including external stimulus information, cognitive information such as plans and errors, and bodily state information (e.g. heart activity, temperature, pain). The fidelity of the feedback monitoring process was measured by Matthews, et al. (1997) using one of the three scales devised by (Sarason, Sarason, Keefe, Hayes, & Shearin, 1986) to operationalise the concept of Cognitive Interference (Sarason, Sarason, & Pierce, 1995). This concept refers to the relative frequency of task-relevant and task-irrelevant thoughts during performance. Cognitive interference increases when task relevant thoughts decline at the expense of increased task-relevant thoughts and the individual has a reduced awareness of task activities.

According to the S-REF model (Wells & Matthews, 1994), when a discrepancy is detected at the lower-level (1) on the basis of external feedback, it is appraised by the S-REF at the second level (2). This act of appraisal will initiate a discrepancy-reduction activity, which affects the sensitivity of the monitoring level (2) to specific stimuli at the lower level (1), i.e. a bias towards discrepancy-relevant stimuli. In addition, activation of the S-REF (via discrepancy-reduction) may involve consultation and elaboration with the upper level (3) of self-knowledge and self-beliefs. The purpose of this liaison is to modify the upper level (3) based on bottom-up feedback from levels (2) and (1). This possibility is indexed by Matthews, et al. (1997) using scales to measure self-focus and self-esteem, derived respectively from Fenigstein, Scheier, & Buss (1975) and Heatherton & Polivy (1991). The existence of persistent discrepancies with respect to external feedback may trigger increased self-focus and in some cases, even threaten self-esteem.

These elements of the S-REF model have been incorporated into the third meta-construct of worry by Matthews, et al. (1997). In this context, worry refers to the act of self-evaluation, e.g. "a re-assessment of personal qualities and goals" (p. 14). It has been hypothesised that the opportunity for self-reflection and worry are determined by task characteristics. For instance, a demanding task may involve sufficient time pressure to suppress self-evaluation activities (Matthews, et al., 1997). This hypothesis was reinforced by empirical data from performance on the working memory task, which indicated no significant association with any aspect of worry. However, both level of self-focus and frequency of task irrelevant thoughts exhibited a significant, negative relationship with successive vigilance performance, i.e. where task activity was sustained and temporal demands were lower.

The relationship between the categories of energetical costs encapsulated in each of the three meta-constructs described by Matthews, et al. (1997) may exhibit different qualitative patterns across different task categories. This approach and analysis is based on the 'state patterning' approach to describe the effects of stress devised by Hockey & Hamilton (1983). Table 3 is adapted from Matthews, et al. (1997) and summarises the influence of different task manipulations on the three meta-constructs of engagement, distress and worry. These data incorporate four studies of working memory, a vigilance experiment, a study of prolonged simulator driving (Desmond & Matthews, 1997) and a study of simulated driving, where control was reduced via the introduction of 'black ice' to induce unavoidable episodes of skidding (Matthews, Sparkes, & Bygrave, 1996). The table illustrates how different qualitative patterns of energetic costs were associated with performance under all four conditions. Short-term, demanding performance on working memory increased both engagement and distress, whereas long-duration vigilance performance depressed engagement. Performance of a prolonged drive had a similar effect to vigilance performance, however subjects also felt distressed. The reduction of control during performance did not influence engagement but increased both distress and worry.

Task type	Engagement	Distress	Worry
Working memory	+	+	-
Vigilance	-	0	0
Prolonged drive	-	+	0
Loss of control (Black ice)	0	+	+

Table 3. The relationship between the three categories of energetical costs and task performance across four experimental scenarios (adapted from Matthews, et al. (1997)). NB: + = increase, 0 = no change, - = decrease.

This conceptualisation of energetical costs indicates the range of internal data that may be incorporated into internal feedback, e.g. awareness of self-directed cognition, physiological symptoms. In addition, there is evidence that categories of costs may show characteristic patterns of association and dissociation due to different task demands. For example, Duval & Wicklund (1972) postulated that self-focus declined during an absorbing task activity, i.e. when engagement increased. Similarly, Desmond, Matthews, & Hancock (1997) reported that declining levels of engagement and rising distress were significantly correlated with physiological symptoms of fatigue, e.g. visual fatigue, malaise, muscular fatigue.

Feedback from internal sources may manifest a range of energetical costs associated with performance (Hockey, 1993; Hockey, 1997). The frequency and magnitude of these costs represent the level of energetical expenditure associated with task performance. These costs may be operationalised in the form of physiological checklists and/or a psychological scales with respect to engagement, distress and worry (Matthews, et al., 1997). It is postulated that increased compensatory costs may be typified as declining engagement and/or rising levels of distress and worry.

2.2.6 The assessment of external and internal feedback

The previous sections reviewed sources of external and internal feedback. Both factors are influential for the formulation of effort policy, i.e. a decision whether to invest mental effort or not.

It is postulated that both internal and external feedback sources are amalgamated into a cognitive-energetical assessment of behavioural efficiency as defined by Schonpflug (1983; 1985). This concept was discussed in Section 1.1. This proposition is based upon two related hypotheses: (a) that effort is a cognitive-energetical entity in the sense that mental effort functions as a cognitive moderator which is energetical in character (Gaillard, 1993), and therefore, (b) effort policy must be based on an appraisal which is sensitive to both cognitive and energetical variables.

An appraisal of cognitive-energetical efficiency requires a context for assessment, which is provided by standards or set-points based on past experience. Therefore, an individual have expectations associated with an appraisal of efficiency, i.e. performance effectiveness x is achieved at energetical cost y. This two-dimensional space provides a flexible rationale for the assessment of behavioural efficiency.

The following table provides a continuum of external/internal appraisals based on feedback from performance and energetical costs (Table 4). This analysis is based on Schonpflug (1985) and those modes of state control proposed by Hockey (1986b). When behaviour is efficient, effective performance is achieved either at reduced or 'standard' energetical cost. If the level of performance effectiveness is matched by the associated level of energetical costs, this is called correspondent efficiency. Inefficient behaviour is defined by those occasions when performance effectiveness is only achieved at accelerated energetical costs.

Performance Effectiveness	Energetical costs	Description	Behavioural Efficiency
+	-	Exceptional performance achieved with a reduction of energetical costs	Best Efficiency
+	0	Exceptional performance achieved at a 'standard' energetical cost	Efficient
0	-	Standard performance achieved in combination with reduced energetical costs	Efficient
+	+	Exceptional performance achieved with high energetical costs	Correspondent
0	0	Standard performance achieved at a standard energetical cost	Correspondent
-	-	Poor performance in combination with reduced energetical costs	Correspondent
0	+	Standard performance is only achieved in combination with increased energetical costs	Inefficient
-	0	Poor performance in combination with standard energetical cost	Inefficient
-	+	Poor performance in combination with increased energetical costs	Worst Inefficiency

Table 4. The assessment of behavioural efficiency using external (performance effectiveness) and internal (energetical costs) feedback in combination. NB: + = positive discrepancy (higher than standard), 0 = no discrepancy (standard achieved), - = negative discrepancy (lower than standard).

A cognitive-energetical monitor is proposed as an executive or supervisory controller (Hockey, 1997) (Figure 7) for feedback appraisal. This monitor receives dual-feedback from external and internal sources. Feedback from overt performance may take the form of overt errors, or a perceptual appraisal of productivity, task failure or the perception of subjective mental workload (Section 2.2.4). Internal feedback is based on CNS activity, physiological symptoms, mood and similar indicators of energetical state (Section 2.2.5). The combination of both sources of feedback yields the two-dimensional appraisal of task efficiency represented in Table 4. This process of self-appraisal is not intended to represent an introspective evaluation of efficiency as a value judgement in its own right, i.e. the individual does not ask himself or herself "how efficient is my behaviour?" The assessment of efficiency is an emergent feature from two distinct sources of feedback.

The most important point to note concerning the appraisal of efficiency is that assessment of feedback is inherently phenomenological, representing the experiential world of the individual. Therefore, an individual must rely on his or her experience and powers of observation, perception and introspection to make consistent and reliable self-assessments.

The process of self-appraisal may be inherently fallible with respect to both external and internal feedback. The subjective appraisal of cognitive-energetical efficiency may be prone to at least three categories of bias or distortion. In the first instance, the appraisal of either internal or external feedback is subject to *attentional selectivity*. For example, Matthews, Carver, & Scheier (1982) pointed out that "selectively attending to aspects of oneself appears to produce effects that are conceptually identical to those produced by selectively attending to aspects of one's environment" (p. 167). In other words, we may attend to internal feedback at the expense of external feedback and vice versa. This particular aspect is a central hypothesis of the S-REF model (Wells, 1994), where it is argued that activation of level (3) concerned with self knowledge via the experience of worry distracts attention from the practical business of task monitoring and performance at level (1) (for detailed explanation, see (Wells, 1994)). The inverse case may occur if a particularly demanding or engrossing biases attention to external information at the expense of internal feedback (Duval & Wicklund, 1972; Matthews, et al., 1997). In this case, neglect of internal feedback may produce substantial fatigue after-effects following task performance (Hockey, 1993; Hockey, 1997), i.e. accumulated energetical costs may only impose themselves on self-appraisal in a retrospective fashion following cessation of task performance.

The second limitation on feedback appraisal is represented by possible *confounding* between internal and external feedback. It is evident from the previous section that increased task difficulty may be associated with rising energetical costs. For example, Carver & Scheier (1990) linked positive and negative affect with respective increases and decreases of task progress. On the other hand, Matthews, et al. (1997) has reported a strong correlation between the subjective assessment of task workload and energetical costs in the form of engagement, distress and worry. Therefore, increased energetical costs, induced via biological or environment stressors tend to amplify the subjective appraisal of task demands and workload (Hancock & Warm, 1989). This link is intuitive, as the net result of declining engagement coupled with rising distress and worry would be to increase the difficulty of task performance. These energetical costs may create a psychological inertia by creating an increasing disinclination to continue performance (Bartley & Chute, 1947), therefore, increased levels of mental effort are required to overcome any 'mental inertia' (Table 4).

The net influence of distress and worry is the proliferation of distracting psychological stimuli linked to reduced confidence, task irrelevant thoughts (Sarason, Sarason, Keefe, Hayes, & Shearin, 1986), negative affect and anxiety-related self-appraisal (Wells, 1994). Therefore, an individual experiencing rising energetical costs must invest a higher level of mental effort in order to compensate for reduced energy and increased distraction, which develops into a vicious circle of spiralling demands on mental effort, i.e. given that future effort investments may be associated with rising energetical costs.

The framework described by Hancock & Warm (1989) provides a conceptual means of reconciling task demands and the demands induced by energetical factors. According to Hancock & Warm (1989), both increased task demand and energetical costs function as sources of input stress on the individual, and the response to either is co-ordinated on an adaptive basis. Therefore, both internal and external factors must be incorporated within a single, cognitive-energetical strategy as effort policy. However, this conceptual unity does not alter the confounding influence of internal on external feedback and vice versa. If the subjective appraisal of either internal or external feedback is very susceptible to the influence of the other, this creates problems for the cognitive-energetical monitor, which is attempting to identify and to adapt effort policy to specific internal and external factors. An example of an internal → external bias was reported by Bartlett (1943), who noted that operators exhibited an increased tendency to blame inadequate apparatus and/or instrumentation for error as a function of increased fatigue. This process of internal/external assessment is particularly problematic as both sources must be integrated in order to estimate efficiency, therefore, tremendous potential exists for a form of 'crosstalk' that may produce inaccurate self-appraisal.

The final category of attentional distortion is linked to the influence of energetical costs on the *appraisal process* itself. The evolution of fatigue/sleepiness has been associated with a range of influences on human information processing (Bartlett, 1953; Bartlett, 1943; Bills, 1931; Dinges & Kribbs, 1991; Easterbrook, 1959; Hockey, 1986a). These attentional effects have included lapses or blocks (i.e. short periods when attentional processing appears to cease), tunnelling (i.e. high selectivity which neglects peripheral or secondary stimuli) and an increased 'indifference range' (i.e. the tolerance of increased levels of error prior to corrective action). The influence of these effects on performance has been well documented. However, as pointed out by Brown (1994), if the capability to perform is jeopardised by energetical factors, then logically, our capability for self-appraisal may be similarly impaired. Therefore, the ability to monitor and respond with respect to either internal or external sources of feedback may be susceptible to lapses, tunnelling and increased indifference. The influence of energetical costs on the process of self-appraisal may lead to flawed estimates of efficiency.

This section has described how external and internal sources of feedback may be reconciled within a self-appraisal of behavioural efficiency. Furthermore, it is argued that emergent efficiency from dual-feedback is the essential concept guiding the subsequent formulation of effort policy. The following section will describe the formulation of effort policy in detail.

2.2.7 Effort policy

The dual appraisal of internal and external feedback is converted into an assessment of cognitive-energetical efficiency, which gives rise to a specific effort policy. This policy may be described with reference to whether effort is invested or conserved.

The decision to invest or to conserve effort was described by Hockey (1986b) as modes of state control. The investment of mental effort represents a generalised policy where mental effort is increased in order to improve or to sustain performance effectiveness, i.e. a "try harder" response. An effort investment policy is the natural response to a perceived increase of task demand or a decline of performance effectiveness. The conservation of mental effort represents the withdrawal of mental effort or the suppression of continued effort investment. In this case, performance quality may be sacrificed to prevent further accumulation of energetical costs such as stress and fatigue. On the other hand, conservation may occur when a desirable standard of performance effectiveness may be attained with lower effort expenditure, e.g. when skill develops, performance quality may be possible with lower effort investment.

The formulation of effort policy is an adaptive process, responding dynamically to sustain task performance and prevent unacceptable levels of stress and/or fatigue. An investment policy will improve or protect performance effectiveness at the expense of energetical costs. A conservation policy performs the opposite function. Therefore, effort policy must be devised on an adaptive basis to compensate for declining performance/rising energetical costs (Hancock, 1986; Hancock & Warm, 1989). This dynamic underlines the importance of self-appraisal with respect to both internal and external sources of information. The framework for effort policy is illustrated below in Figure 9. This framework is based on an analysis of workload and performance by (De Waard, 1996) and the model of sustained attention and stress produced by Hancock & Warm (1989).

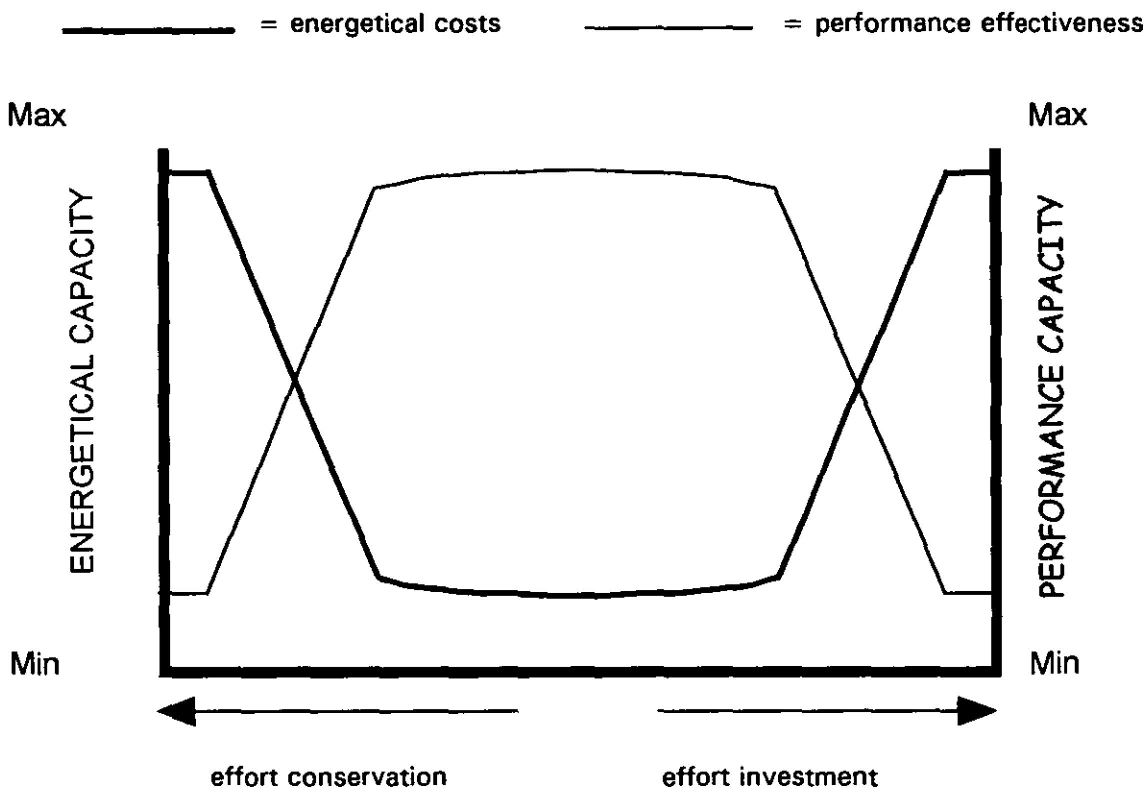


Figure 9. A framework for mental effort policy based on analyses by De Waard (1996) and Hancock & Warm (1989).

According to this framework, a stable equilibrium is apparent when energetical costs are minimal and performance is maximal at the centre of the horizontal continuum. Therefore, the individual may dynamically formulate effort policy in order to remain as close to the central region as possible. This may be accomplished by alternating periods of effort investment and conservation, to allow those energetical costs, which may have accumulated during effortful periods to dissipate during periods of conservation. At both edges of the horizontal continuum, poor performance effectiveness co-exists alongside maximum energetical costs. This undesirable situation may arise due to excessive investment (on the right-hand side of the diagram) or unrestrained conservation (on the left side).

In both cases, the formulation of effort policy occurs within finite limits determined by: (a) capability for performance effectiveness, and (b) energetical capacity. The former is determined by individuals' level of performance efficiency. This concept was originally proposed by Eysenck (1982) and operationalised by Meijman (1995) to represent the relationship between effort (energetical input) and performance quality (overt output). Efficient individuals attain good performance at low effort investment, whereas less efficient individuals must expend greater effort to attain the same level of good performance. This capacity for performance may be determined by an individual's level of skill, which in turn, sets upper and lower limits on performance effectiveness.

Hockey (1993; 1997) described a number of latent decrements which may occur prior to primary performance breakdown. These decrements represent strategies or heuristics to enhance performance efficiency without sacrificing performance quality. The successful deployment of latent decrements may extend the lower limit of performance capacity. One response to declining performance capacity is termed *subsidiary task failure*. This decrement is similar to attentional tunnelling (Easterbrook, 1959), where performance concerning peripheral perceptual stimuli or subsidiary task elements is neglected in order to protect performance involving central stimuli or crucial task elements. This decrement represents an adaptive attentional strategy to preserve performance in the face of either high task demands or extreme energetical costs. The second decrement is termed *strategic adjustment*, which is described as a within-task shift to simplistic task strategies. Therefore, the individual may adapt task strategy to a routinised, regular sequence that may be inefficient but is relatively robust and resistant to errors (Sperandio, 1978). Alternatively, if task pacing can be controlled, the individual may wish to slow the task. This decrement has also been observed in experimental subjects working in the presence of noise (Schonpflug, 1983). The third decrement is termed *fatigue after-effects* and refers to a preference for low-effort strategies, which may follow a period of high task demand or an exhausting work activity. This decrement represents a period of effort conservation following a work period that has incurred high energetical costs, and therefore, may correspond to an attempt to restore equilibrium (Figure 9).

The latter concept of energetical capacity is more complex, being equivalent to the level of remaining effort reserves and the individuals' tolerance for discomfort (Pribram, 1980), i.e. reduced effort reserves = increased discomfort. This capacity for discomfort may be determined by goal characteristics (e.g. commitment to the task goal) and individual traits, i.e. Matthews, Schwean, Campbell, Saklofske, & Mohamed (in press) provide a review and synthesis of personality literature, which indicates that neurotic individuals may be more sensitive to energetical costs than non-neurotics.

The framework for mental effort investment illustrated in Figure 9 is subject to limitations due to respective capacity of the individual to adapt effort policy to external and internal feedback. This adaptive capability (Hancock & Warm, 1989) is limited by the level of effort reserves available to the individual. However, a number of other limitations may be considered. It is postulated that effort policy may be executed in a fashion that is either efficient, correspondent or inefficient (Table 4).

A correspondent investment policy would improve performance effectiveness and increase energetical costs, whereas efficient investment results in the same improvement of performance without any subsequent increase of energetical costs. On the other hand, no influence on performance or declining performance coupled with increased costs would indicate an inefficient effort investment policy. The same logic may be applied to effort conservation policies, i.e. an efficient conservation policy would reduce energetical costs without affecting performance effectiveness.

It is proposed that the limitations of effort policy are synonymous with the limitations on performance efficiency. Inefficient performance will produce a greater number of errors and/or high costs at a faster rate per unit time than efficient task performance. The linkage between effort policy, efficiency and limitations is considered within the adaptive framework originally outlined by (Hancock & Warm, 1989) and illustrated in Figure 10 below.

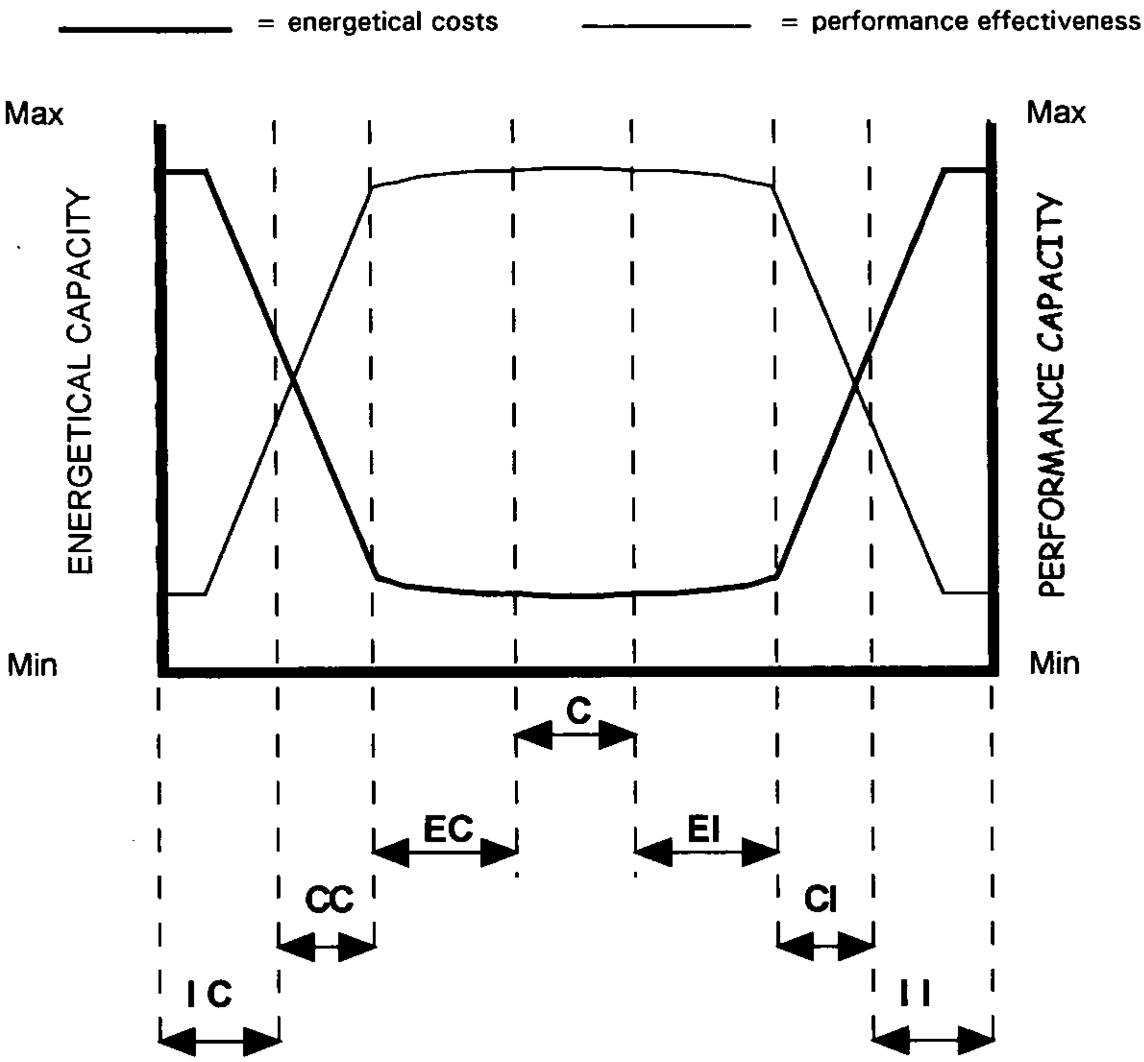


Figure 10. Appended version of effort policy framework. Abbreviations: C = comfort zone (based on Hancock & Warm (1989)), EC/EI = efficient conservation/investment, CC/CI = correspondent conservation/investment, IC/II = inefficient conservation/investment.

The use of the comfort zone to indicate complete equilibrium is taken from the model devised by Hancock & Warm (1989). This zone indicates an absence of either investment or conservation in combination with minimum energetical costs and maximum performance. This zone represents an ideal that may not be possible to sustain for a significant period. The movement to the left and right of the horizontal continuum moves through zones associated with efficiency, correspondence and inefficiency. When a policy is efficient, it is achieving a primary goal without incurring costs or degraded performance. An inefficient effort policy fails because it does not achieve a primary goal and still incurs costs or degraded performance. Once this point has been reached, the individual is at the limits of their performance capability and tolerance for discomfort.

Mental effort policy may be described in terms of investment or conservation. The primary goal of the former is to improve performance effectiveness, whereas the latter is concerned with the reduction of energetical costs. The adaptive, dynamic formulation of mental effort policy is based around the antagonism between investment and conservation. Both mental effort investment and conservation may be assessed in terms of efficiency.

2.3 A model of mental effort regulation

A model of mental effort regulation is proposed based on the concepts and literature reviewed in the previous sections (Figure 11). This model is hierarchical and based on the two-tier framework outlined by Broadbent (1971) and Hockey (1997) (Section 2.2.1).

This model was devised to investigate mental effort regulation under specific circumstances. It is assumed that the individual is an experienced operator and capable of acceptable task performance, i.e. the model does not consider the acquisition of skill. It is also assumed that the individual is performing the task over a sustained period, e.g. 30min or above, and therefore, the exhaustion of finite effort reserves is a distinct possibility.

The cognitive-energetical monitor represents the upper-level, supervisory controller, capable of monitoring and regulating mental effort. The purpose of this mechanism is to avoid catastrophic task failure and unacceptable levels of discomfort due to stress and fatigue. In terms of the effort economy described in Section 1.2, the goal of the monitor is to avoid the collapse of performance and the exhaustion of finite mental effort reserves.

This task of monitoring and controlling is achieved via dual-feedback from internal and external sources. Those internal symptoms of sleepiness, physical discomfort, fatigue, distress and worry originate from the energetical state (Section 2.2.5). This concept represents the internal psycho-physiological status of the individual based on physiological symptoms, CNS activity and mood. This information is supplemented by feedback originating from the impact of overt performance in the external world (Section 2.2.4). This feedback is composed of an appraisal of performance effectiveness based on the frequency of errors and the perception of subjective mental workload. An amalgamation of internal and external feedback takes is performed within the cognitive-energetical monitor (Figure 11).

The assessment of internal and external feedback yields a two-dimensional continuum of high/low energetical costs and good/bad performance (Table 4). In general, increased errors lead to the formulation of an effort investment policy. This policy will increase the proportion of controlled processing and improve the sensitivity and precision of perceptual-motor behaviour, i.e. by increasing the speed/accuracy of target detection, increasing the frequency of attentional checks during response preparation. In addition, effort investment may be associated with a degree of sympathetic activation, which may give rise to energetical costs. If effort were invested continuously for a sustained period, the resulting accumulation of energetical costs may induce an experience of fatigue, stress or a combination of both for the individual via the internal feedback loop. In this case, the individual may reduce the level of effort investment in order to alleviate the discomfort associated with energetical costs. This strategy is known as switching to a policy of effort conservation. It is presumed that a reduction of effort investment will at least halt the continued accumulation of energetical costs.

There is a degree of mutual dependence between the assessment of external and internal feedback. If error rate is high, the individual may experience task-related stress due to perception of task failure. Hence, an appraisal of external feedback may exacerbate the accumulation of energetical costs. A similar phenomenon is observed when changes in energetical status impinge on task performance. For example, a sleep-deprived individual may experience lapses in concentration that provoke an increased frequency of task errors. Similarly, the experience of stress states such as distress and worry may distract attention from the task at hand, leading to an increased frequency of errors due to cognitive interference. This inter-dependence means that effort must be invested to protect performance from the debilitating influence of energetical costs. Therefore, rising energetical costs may be ambivalent with respect to their influence on effort policy and much depends on feedback from external sources.

It is hypothesised that the appraisal of internal and external feedback is a perceptual process and subject to a degree of inherent bias. These sources of bias may be described as attentional or energetical in nature. For example, a highly demanding task may focus attention on the external feedback loop at the expense of the internal assessment of energetical costs. On the other hand, the presence of energetical stressors such as sleep deprivation and alcohol may influence the appraisal of external and internal feedback. Certain categories of stressor variables such as sleep deprivation may emphasise the inherent discomfort of sustained performance. Whereas chemical agents such as alcohol may increase well-being and reduce awareness of errors with consequences for the formulation of effort policy.

Once an effort policy is formulated, several outcomes are possible. If an investment policy is successful, overt performance should improve or stabilise and energetical costs may rise. If investment is efficient, improved or stable performance may be achieved in conjunction with reduced energetical costs. (Figure 10). A successful policy of effort conservation would result in reduced energetical costs. In this case, the goal of the monitor is to alleviate the energetical discomfort of continued task activity. In both cases of investment and conservation, the ideal for the cognitive-energetical monitor is to maximise performance and minimise energetical costs, i.e. the comfort zone in Figure 10. The obverse case of unsuccessful and/or inefficient effort policy is apparent when investment fails to improve or stabilise performance effectiveness, or when conservation has no impact on energetical costs.

This process of feedback appraisal and policy formulation is highly dynamic and may fluctuate on a minute-by-minute basis. The feedback from energetical state and overt performance represents the success or failure of the current effort policy. Hence, inefficient or failed effort policies manifest themselves via feedback to the cognitive-energetical monitor, which adapts effort policy appropriately.

It has been argued that the process of mental effort regulation is perceptual and prone to a variety of attentional biases and distortions (Section 2.2.6). These factors may lead to the inappropriate formulation of effort policy. For example, a workaholic may attend to external feedback at the expense of internal feedback, hence risking the threat of stress-related illness. Similarly, a hypochondriac may monitor internal symptoms and sensations to the extent that events in the external world pass unnoticed. In addition, it may be difficult for the monitor to establish causality due to confounding between internal and external signals. It is likely that certain energetical states such as extreme sleepiness or exposure to drugs are capable of degrading perceptual-motor behaviour and subsequent performance. In this case, the monitor will receive feedback of high costs coupled with poor performance, where the latter is a direct result of the former. Similarly, the experience of repeated errors or task failure may provoke a stress reaction from an individual leading to an identical scenario. Therefore, the monitor may encounter significant difficulties when formulating effort policy under certain circumstances.

The proposed model of mental effort is illustrated below in Figure 11. The cognitive-energetical monitor receives feedback from internal and external sources. This process of dual-appraisal is used to formulate effort policy, i.e. whether to invest or conserve mental effort. An investment of mental effort may lead to increased energetical costs, hence the dashed line in Figure 11. It is assumed that effort investment influences perceptual-motor performance directly and overt performance indirectly. The enactment of an effort policy will influence both overt performance and the internal, energetical state and hence, the feedback loop will begin once more. It should be noted that both internal and external feedback is subject to a degree of mutual interference. High error rates during overt performance may induce energetical changes due to the experience of distress and worry. It is also hypothesised that extreme energetical costs (such as fatigue and stress) may directly influence perceptual-motor behaviour with consequences for task performance. This inter-dependence is illustrated via the grey feedback loops at the bottom of Figure 11.

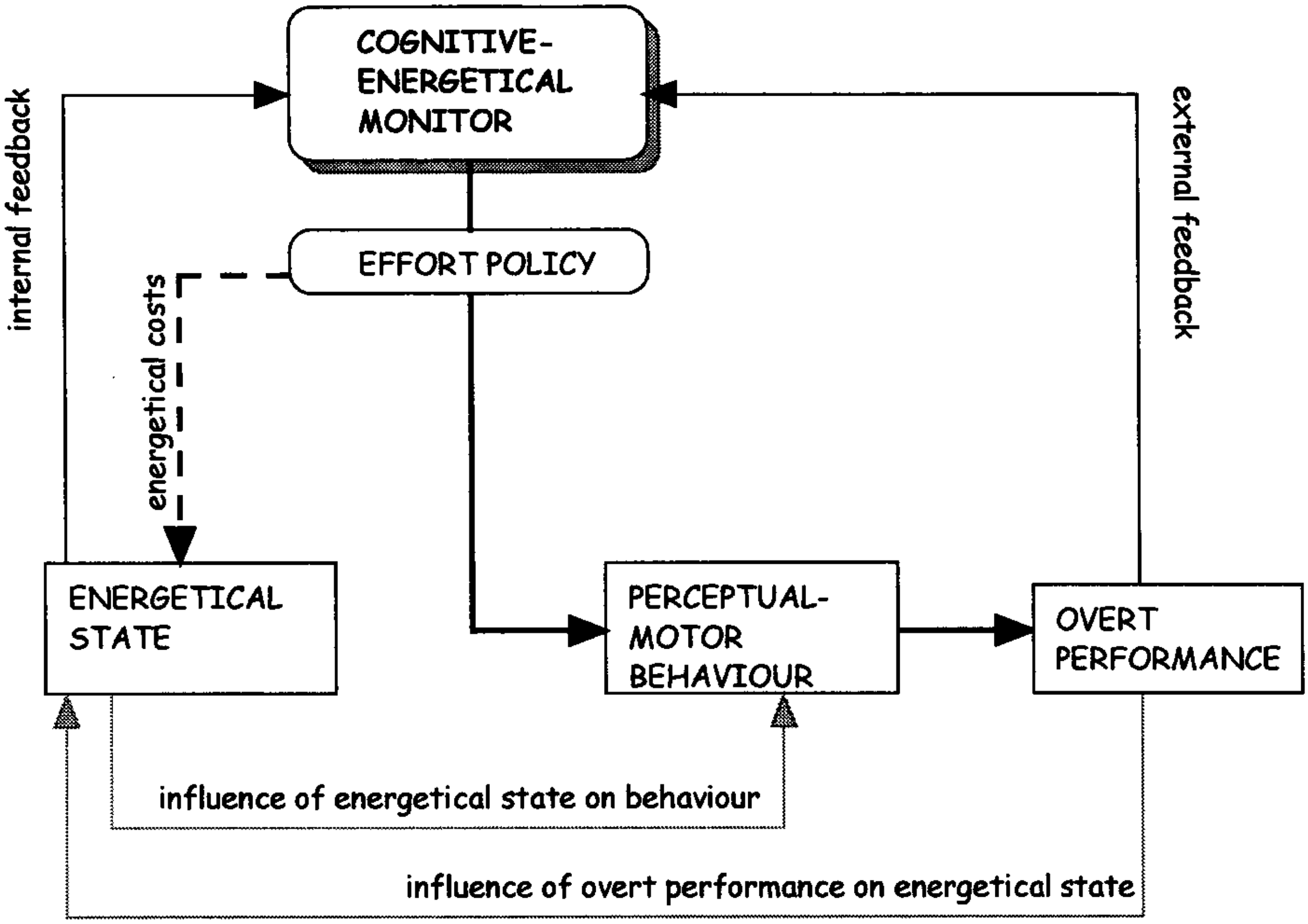


Figure 11. The thesis model of mental effort regulation

3 INTRODUCTION TO EMPIRICAL WORK

3.1 Hypotheses arising from the model of effort regulation

The previous chapter concluded with a description of a model of mental effort regulation. This model represents the basis of the empirical programme described in the following four chapters. The central predictions encapsulated within the effort regulation model are as follows:

1. The decision whether or not to invest mental effort is based on a cognitive-energetical appraisal from two sources of feedback: (a) performance cues from the external world and (b) internal cues regarding psycho-physiological, energetical state.

2. The individual may adopt one of three effort strategies based on an assessment of external errors and internal costs: (a) effort may be invested, (b) effort may be conserved or withdrawn and (c) effort may be sustained.

3. Competent performance is achieved by higher levels of effort investment and rational effort investment, based on an awareness of both external and internal feedback.

4. The presence of stressors detrimental to sustained performance, such as sleep deprivation, may require effort investment to protect task performance from the influence of energetical costs.

5. The presence of chemical stressors, such as alcohol, may influence the process of appraisal of external and internal feedback.

6. The presence of performance feedback will highlight the salience of external feedback and lead to a tendency for effort investment.

7. High task demands emphasise the salience of the external feedback loop and the subsequent regulation of mental effort.

Each experiment will emphasise a different aspect of these effort hypotheses. In addition, several experimental tasks and testing environments have been adopted for the study of mental effort. This approach to the research is described in the following section.

3.2 Test environments and experimental tasks

The thesis has focused on sustained perceptual-motor performance as a broad category of task activity. Two different categories of perceptual-motor task are used during the empirical programme described in the following chapters. The first was a computer-based task that incorporated a zero-order, lateral tracking activity with a visually-based, cognitive/successive, vigilance task (Warm & Dember, 1998). This type of task was used in the studies described in Chapters 7 and 4. The latter studies are representative of applied research using a simulated driving task as a perceptual-motor task. The experiment described in Chapter 5 employed a fixed-base driving simulator (i.e. stationary vehicle with original controls, computer-generated driving scene projected onto a large screen), whereas the investigation in Chapter 6 utilised a modest approximation of the driving task (i.e. a PC-based driving task in conjunction with a steering wheel console).

Both testing environments and tasks were employed to fulfil different research requirements. The use of a laboratory-based tracking task permits tight control over important experimental variables, despite the nature of the task being essentially meaningless and having limited correspondence to real-world activities. On the other hand, the employment of a simulated driving environment represents an attempt to duplicate a real-world activity within an artificial setting. The driving simulator task is a complex task environment relative to the tracking task, associated with a greater range of experimental variables. In addition, the subjects have greater control over performance in this simulated environment and behaviour is at least partially determined by prior training and experience with an everyday task. The increased range of task variables in conjunction with higher subject control and experience may conspire to reduce the degree of control over behaviour wielded by the experimenter. This antagonistic relationship between ecological validity and experimental control across testing environments was described by Parkes, Fairclough, & Ross (1991).

Both task categories have a long research history with respect to dynamic task performance and the study of energetical variables such as fatigue. Literature reviews on tracking tasks for the study of control dynamics have been performed by Hammerton (1981) and Wickens (1986a). The use of laboratory tasks and other driving simulations for the investigation of driver fatigue was reviewed by

Haworth, Triggs, & Grey (1988).

Both types of task share a number of common features. For example, both are concerned with manual control and lateral movement (i.e. tracking, steering). Both tasks also include an implicit response time component, i.e. target detection, braking to obstacles or other vehicles. In addition, it could be reasonably argued that both tasks conform to a similar profile according to the multiple resource framework (Wickens, 1980; Wickens, 1984), i.e. both task are visual, involve processing of spatial data and demand a manual response.

However, there are areas of substantial divergence between tracking and driving behaviour. Firstly, subjects may manipulate speed and task pacing during the latter activity, whereas no such facility was available during the tracking task. In addition, tracking activity may be classified as two distinct types: pursuit tracking (i.e. replicating the speed/direction of a target) or compensatory tracking (i.e. detecting the deviation of a cursor from a target position and making an appropriate correction) (Hammerton, 1981). The laboratory task used in this experimental research was designed as a pursuit tracking task, whereas driving behaviour may be characterised as principally compensatory in nature, i.e. we correct for dynamic deviations from lane boundaries. It should be added that control dynamics differ substantially between the tracking task and simulated driving (Wickens, 1986a). The former is characterised as zero-order control, i.e. the magnitude and direction of the input device movement are translated directly to cursor movement onscreen. On the other hand, simulated driving involves steering wheel input that contains an additional velocity component, i.e. cursor movement is a product of steering direction, magnitude and velocity. This type of control dynamic is termed first-order tracking.

A large amount of research has been conducted into the links between psychological impairment (due to alcohol, drugs and fatigue) and driving safety. For ethical reasons, a substantive proportion of this research is not conducted in the field on the real road. Researchers must rely on laboratory tasks, driving simulators and closed-course facilities in order to test performance within acceptable ethical constraints. This climate has fostered a degree of methodological conflict between individual researchers. On one hand, laboratory-based researchers believe that driving behaviour is too complex to yield consistent data in the field. Therefore, they propose to decompose the driving task into its constituent components (e.g. visual search, manual control, decision-making) and to investigate each component in isolation within a laboratory setting, e.g. (Clayton, 1980). The strategy represents the application of reductionism as a means of establishing a high degree of experimental control. Those in opposition object to this approach on the grounds of ecological validity. They argue that sub-components of the driving task are not enacted in isolation, but as a complex act of co-ordinated control. Therefore, these researchers believe that studies of driving behaviour should aim for realism and representativeness. This approach begs a number of ethical questions if taken to an extreme. For the most part, these researchers argue that simulators and closed-courses represent a compromise between ecological validity and ethical investigation, e.g. (Gawron & Ranney, 1988). This methodological issue was addressed and discussed in more detail by Sanders (1985).

The methodology adopted during the current research has attempted to study mental effort with reference to an artificial laboratory task and a simulation of a real task. The use of mixed test environments may have a number of advantages over the exclusive use of one or the other, namely: (a) it is possible to identify key variables with greater clarity within the less complex laboratory setting, (b) it is possible to manipulate variables with greater precision in the laboratory setting, i.e. because task-pacing is computer-controlled, (c) the usage of laboratory and applied testing environments provides an indication of the ecological validity of hypotheses under investigation, and (d) the inclusion of real-world skills permits an assessment of the relevance of hypotheses for well-learned tasks, i.e. a contrast between performance within a novel domain and performance in a real-world domain, where subjects bring real-world experience and training to the study.

The current document is concerned with mental effort regulation as basic theoretical research and as applied topic with real-world relevance. It is anticipated that the inclusion of different test environments will reinforce the relevance of effort regulation in both domains.

3.3 The measurement of mental effort

The quantification of mental effort is described with reference to measures of brain metabolism, hormonal activity and psychophysiological function. With respect to the empirical chapters, it was only possible to quantify mental effort via psychophysiological activity. However, a brief review of all three categories of measurement is warranted for the sake of completion.

3.3.1 Brain metabolism and mental effort

Brain metabolism is the logical candidate to provide a quantification of mental effort. The measurement of cerebral metabolism represents the fundamental substrate of effortful activity. This is both a strength and a weakness, as cerebral metabolism is the fundamental substrate of psychological activity. These techniques were not employed in the empirical studies, however, a brief review is warranted for the sake of completion.

As stated earlier, the energy storage capacity of the human brain is very limited, therefore glucose is supplied to the brain from the blood on a continuous basis. The major location for glucose metabolism is at the synapse, and Positron Emission Topography (PET) techniques may be used to index the cerebral metabolic rate for glucose (CMRglu). The principles of PET measurement are based on the fact that biologically active elements, e.g. fluorine, oxygen, carbon, have neutron-deficient radioisotopes. These radionuclides have a limited half-life (i.e. around 110 minutes) during they decay by the emission of positrons. Using these elements as markers in combination with appropriate scanning apparatus, it is possible to quantify the rate of brain metabolism across different areas of the brain (Pawlik & Heiss, 1989).

Cohen, et al. (1988) performed a study of the localisation of CMRglu rates within the brain during a sustained attention task. These authors performed an independent samples design where the PET data from a "resting" group (N=9) with a group who performed a sustained, auditory discrimination task (N=27). This task involved the subject listening to a series of 1.0s duration tones, presented at three levels of intensity and having to press a button when the lowest volume tone was detected. Subjects performed this task for 35 minutes. Contrary to earlier findings (i.e. Reivich, Alavi, Gur, & Greenberg (1985)), this study revealed only regional differences in the prefrontal cortex due to sustained attention (i.e. global CMRglu rates were not significantly different). These authors attempted to correlate the number of hits and false alarms with regional CMRglu rates. Negative correlations between false alarms and CMRglu rates were noted in the medial, left anterior and right anterior frontal cortex.

Haier, et al. (1988) performed a similar study, contrasting PET data for three independent subject groups over three activities: an abstract reasoning task, a visual vigilance task and a control group (who received only non-target stimuli from the second task). The visual vigilance task consisted of a 30 minute vigil where subjects were exposed to degraded exposures of single digits (one every 2.0s) and instructed to respond to the digit 0. The abstract reasoning task was conducted over a similar duration. The data revealed a significant increase of CMRglu rate in the right hemisphere for both experimental tasks, however, no particular area was implicated. An examination of the relationship between cerebral metabolism and performance revealed a negative correlation was found for CMRglu rates and performance on the abstract reasoning task (i.e. poor performance was associated with high levels of CMRglu rate). Whilst correlations between sensitivity (d') and CMRglu rates in the right, inferior, temporal lobe were positive, i.e. higher cerebral glucose metabolism was associated with accurate performance.

Techniques to measure cerebral blood flow (rCRF) were originally devised to map brain vascular physiology, rather than to index neuropsychological function. Nevertheless, there is a strong association between rCBF and energy metabolism in the brain (Pawlik & Heiss, 1989).

A recent pioneering study into the relationship between cerebral blood flow and the vigilance decrement was conducted by Mayleben, et al. (1998). These authors used transcranial Doppler sonography (TCD) to monitor the velocity of cerebral blood flow in the left and right hemispheres during two categories of vigilance task. The first type of task involved SUCCESSIVE vigilance where subjects were asked to make an absolute judgement regarding the presence of exceptional target items, based on their memory of the normative stimulus characteristics, i.e. subjects presented with two lines and asked to respond when both were 3mm taller than *usual*. The second SIMULTANEOUS vigilance task involved a comparative judgement, where subjects were presented with two lines and asked to respond when one line was 2mm taller than the other. Warm & Dember (1998) have provided evidence that successive tasks are inherently more demanding than their simultaneous counterparts due to a working memory component. Ten participants performed 30min vigils of both types of vigilance task, the stimuli were monitored at a high rate of 30 events per minute and 3% of the events contained targets.

The results of this study are illustrated below in Figure 12. As expected, SUCCESSIVE vigilance tasks produced a higher vigilance decrement over time compared to the simultaneous tasks. However, these performance data were mirrored by changes in the velocity of cerebral blood flow. Furthermore, the velocity of cerebral blood flow was higher in the right hemisphere than the left hemisphere during the SUCCESSIVE vigilance task only.

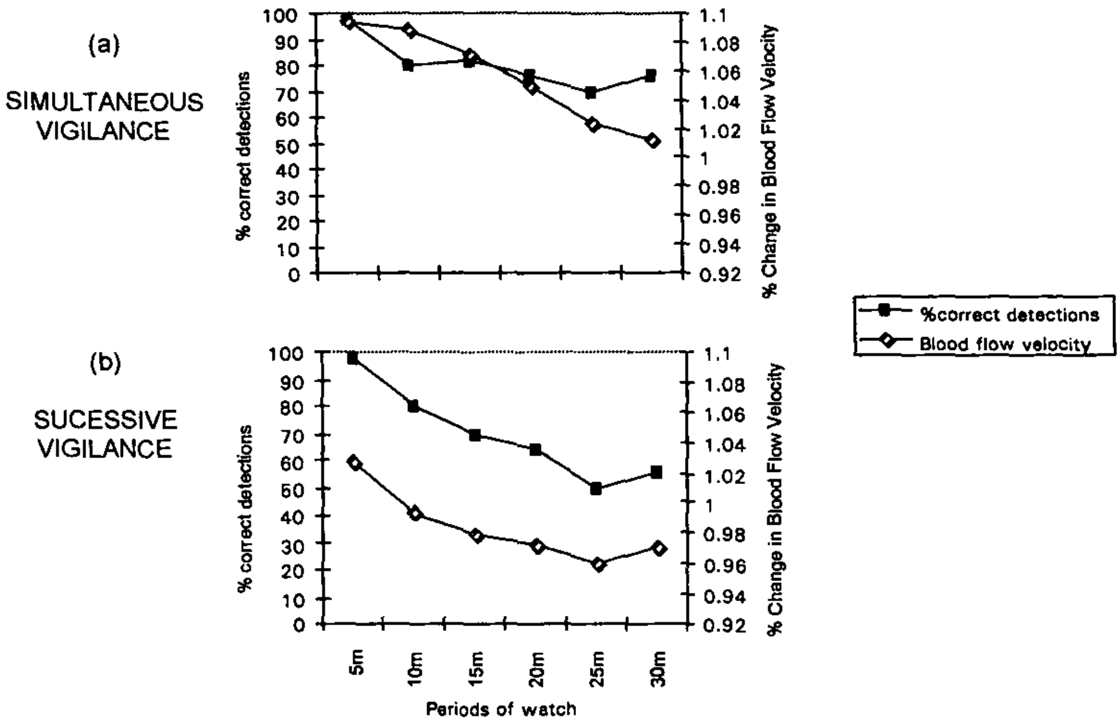


Figure 12. Performance data and mean cerebral blood flow velocity (right, middle cerebral artery) for both sucessive and simultaneous vigilance tasks across a 30 minute vigil (Mayleben, et al., 1998).

These data illustrate a correlational relationship between performance and cerebral metabolism. In addition, the relationship over time and task demands appears compatible with a fuel metaphor for mental effort, i.e. increased metabolism in response to increased task demands, decline of metabolism with time-on-task.

3.3.2 Hormonal changes associated with mental effort

Data on hormonal or endocrine changes are usually collected via urinanalysis. Rather than blood sampling (which is rarely used and difficult to implement), urinanalysis techniques rely on the metabolites of hormonal activity to index endocrine changes (Benton, 1987).

Frankenhaeuser and her colleagues have conducted a programme of research into the relationship between hormonal excretion and human performance for over two decades (Frankenhaeuser, 1980). Frankenhaeuser and her colleagues performed various experimental manipulations to assess the catecholamine response to conditions of "overstimulation" and "understimulation." Their hypotheses were based on the assumption that both extreme boredom and extreme excitement created a disturbance of the individual's cognitive-emotional equilibrium. The subjective response to this disturbance was deemed to be characterised by increased stress and an investment of effort (in order to correct the disturbance).

For example, in one study (Frankenhaeuser, Nordheden, Myrsten, & Post, 1971) 28 subjects paticipated n three experimental sessions: (a) a sensory vigilance task where the subject had to discriminate an increase of light intensity, (b) a multitasking task involving manual responses to concurrent visual and auditory reaction time tasks and (c) a control session where the subjects sat reading magazines. All sessions lasted for 3 hours. Data analysis revealed an increase of catecholamine release for both experimental tasks compared to the control session. Levels of adrenaline showed a slight increase during the vigilance task and a pronounced decline during multitasking, whereas noradrenaline declined during vigilance and showed an increase when the subject was multitasking (all catecholamine data was indexed by urinanalysis). In addition, Frankenhaeuser, Nordheden, Myrsten, & Post (1970)discovered that those subjects with the highest scores on the vigilance task showed elevated levels of adrenaline, whereas lower execretion of adrenaline was associated with superior performance during the multitasking condition. Therefore, it would appear that adrenaline may modulate performance when the individual is "understimulated", whereas noradrenalin exerts a influence in the "overstimulation" situation.

In a later study, Frankenhaeuser and Johannson (1971) studied performance during a variant of the Stroop task. In this study, subjects were asked to respond to both a "single-conflict" (i.e. conventional Stroop colour-word interference) and a "double-conflict" (i.e. the names of colours were presented to the subject via auditory means during Stroop presentations). These authors noted that the more effortful "double-conflict" produced a pattern of reduced performance combined with elevated adrenaline excretion and subjective distress.

A number of studies have been performed to examine the role of catecholamine excretion during performance under environmental stressors. For example, Lundberg & Frankenhaeuser (1978) studied the ability to perform mental arithmetic under different noise conditions. In this study, subjects either experienced high or low levels of personal control over noise intensity, i.e. subjects allowed to set noise intensity themselves (high control) or had the level of intensity set by a 'yoked' partner (low control). Urinanalysis revealed elevated levels of noradrenaline and cortisol excretion during the low control condition. No differences in adrenaline excretion were found.

The issue of personal control obviously had a large impact on subjects' perceptions and performance during the noise study. The primacy of personal control was highlighted by two later studies designed to induce low and high control situations. In the low-control study (Lundberg & Forsman, 1979), subjects performed a sensory vigilance task involving high monotony and unpredictability. The high-control experiment (Frankenhaeuser, Lundberg, & Forsman, 1980) involved a choice-reaction time task, where subjects had the opportunity to choose and modify the stimulus rate (in order to optimise performance) throughout the experimental session. In both cases, self-rated distress and effort was compared with cortisol and adrenaline excretion. Data from both studies are illustrated in Figure 13 below.

It was noticeable that conditions of low control promote an elevation of both subjective effort and distress. Therefore, subjects were invested effort under conditions of discomfort. This pattern of behaviour was associated with a combined elevation of cortisol and adrenaline. When subjects were provided with a high degree of personal control, this pattern was altered. In terms of hormonal activity, adrenaline excretion increased whilst cortisol levsls were reduced. Frankenhaeuser (1987) described the former "effort with distress" pattern as characterising a stressful working situation. However, the "effort without distress" pattern which typifies task involvement in a positive fashion is characterised in the high control situation. The positive experience of "effort without distress" is evident by the subjective ratings where effort is higher than the low control situation, but is accompanied by a decline of distress.

The consequences of elevated cortisol for subjective stress have been confirmed elsewhere. For example, in the review by Wesnes & Warbuton (1983) evidence was presented that corticosteroid elevation is intimately associated with the degree of subjective uncertainty experienced by the individual, i.e. as provoked by real-life emergency situations. In addition, the injection of corticosteroids into subjects produces a tense, alert state marked by irritability and emotional lability.

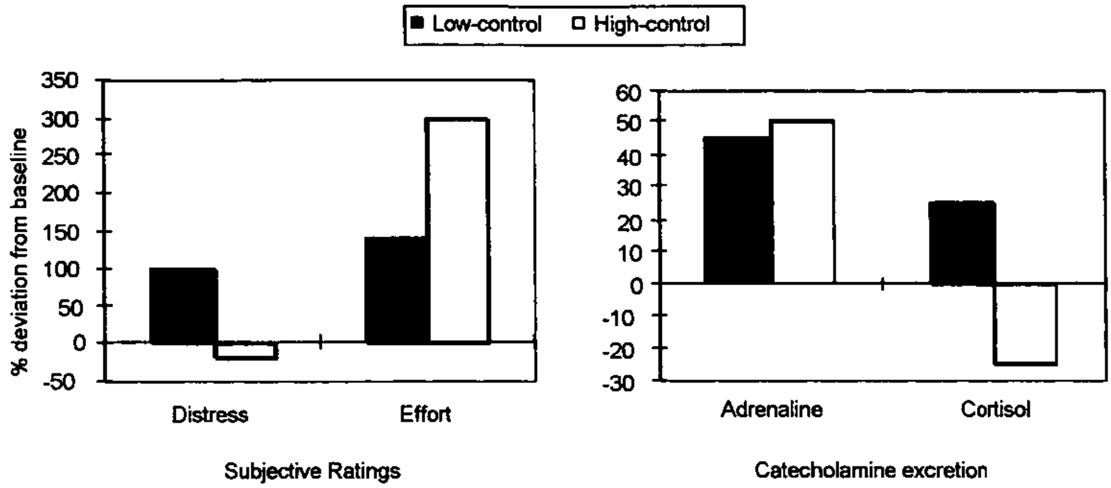


Figure 13. Subjective ratings and catecholamine excretion under conditions of low (Lundberg & Forsman, 1979) and high control (Frankenhaeuser, 1987).

These data illustrate three important points: (1) that adrenal-medullary excretion is associated with task involvement whilst adrenal-cortical excretion is linked to distress, (2) that not all experiences of effortful activity or fuel consumption are perceived in a negative fashion by the subject, and (3) that the appraisal of the task situation is crucial factor which determines the interaction between adrenal-medullary and adrenal-cortical activity. This interaction determines whether "cognitive fuel" is consumed under negative or positive conditions.

The investment of mental effort is perceived to be synonymous with controlled processing described by Schneider & Shiffrin (1977), and (Shiffrin & Schneider, 1977). Paus, Mates, Radil, Hampl, & Husek (1988) performed a replication of their basic paradigm, where a consistent mapping task (CM) involved detecting letters from an array of digits, contrasted with a variable mapping task (VM) where random digits were located from an array of digits. Nine subjects participated in the experiment, performing a 60min vigil under both CM and VM conditions. As expected, the number of hits in the CM condition was significantly higher than the VM condition. In addition, urinanalysis revealed elevated levels of noradrenalin and dopamine during the VM task (involving controlled processing and mental effort). The elevation of dopamine is suggestive given its proposed linkage to motor action control. No equivalent effect was found for either adrenaline or cortisol.

3.3.3 Psychophysiological measures of mental effort

Kahneman (1973) dealt with the problem via the proposition that any prospective physiological index of processing load (i.e. mental effort) should fulfil three criteria to distinguish itself from miscellaneous sources of arousal, these criteria were as follows: (a) be sensitive to within-task variations in task demands, (b) reflect between-task differences elicited by qualitatively different cognitive operations, and (c) should capture individual differences in processing as individuals of different abilities perform the same cognitive task.

Through extensive research and review, Kahneman (1973) selected pupillary dilation as the best candidate measure with respect to all three criteria. He reported a number of studies using a digit transformation task in conjunction with pupillary measurement (Kahneman, 1973). Using a number of manipulations, Kahneman (1973) demonstrated that pupillary dilation increased steadily during digit presentation (i.e. assuming task demands are additive with each sucessive digit).

This work was fortified by Beatty (1982a) who studied the relationship between auditory vigilance performance and pupillary response. His analysis distinguished between two components of the pupillary response: (a) a phasic, pupillary dilation in response to targets and non-targets, and (b) the tonic or baseline pupillary diameter. Beatty (1982a) found the archetypal vigilance decrement (i.e. reduced sensitivity, conservative shift in response criterion) over 48 minutes of monitoring. These changes in performance were mirrored in the pattern of tonic pupillary diameter, but not in the baseline pupillary diameter. Beatty (1982b) reviews these findings amongst other data to endorse pupillary dilation as a measure of mental effort in line with the three criteria proposed by Kahneman (1973).

A second candidate psychophysiological candidate to index mental effort (as opposed to arousal) was proposed by Mulder & Van Der Meulen (1973). These authors suggested that heart rate variability (indexed as the interbeat interval (IBI) of the heart in the time domain) may be a suitable candidate measure. However, further work suggested that raw IBI variability was too global a measure (i.e. susceptible to general activity in the CNS) and in a series of papers (Mulder, 1979a; Mulder, 1979b; Mulder, 1985; Mulder & Mulder, 1981), Mulder and colleagues suggested a two-step process whereby the raw IBI data is treated via power spectrum analysis to yield:

(a) a *low-frequency band* (0.02-0.06Hz) where energy represents vasomotor activity involved in the regulation of body temperature

(b) a *mid-frequency band* (0.07-0.14Hz) where energy represents the short-term regulation of arterial blood pressure

(c) a *high-frequency band* (0.15-0.50Hz) where energy reflects the effects of respiratory activity on IBI regularity.

These authors have claimed that the mid-frequency band (also known as the 0.1Hz sinus arrhymia) is sensitive to the influence of mental effort. The conception of mental effort used by these researchers differs slightly from the one employed by Kahneman (1973). However, Mulder (1986) proposed that mental effort increased if task demands rise (as claimed by (Kahneman, 1973)) or if performance must be maintained in the presence of stressors such as noise and sleeplessness.

The 0.1Hz measure of sinus arrhymia has demonstrated sensitivity to a number of dimensions concerning stimulus evaluation, these have included: (a) the number of dimensions in a multidimensional classification task, (b) the number of comparisons in a sentence comprehension task, and (c) the number of items to be retained in a continuous working memory task (Mulder, 1979a; Mulder, 1979b; Mulder, 1985; Mulder & Mulder, 1981). These findings were replicated by Van Dellen, Aasman, Mulder, & Mulder (1985) and Aasman, Mulder, & Mulder (1987) who also employed a continuous working memory task (i.e. subjects asked to memorise a set of target items to be identified during subsequent presentation) and found that target loads of 4 or more caused power in the mid-frequency bandwidth to significantly decline.

It has been suggested that the sensitivity of the 0.1Hz component outside of these memory manipulations may be rather limited (Jorna, 1992). However, several authors have successfully used the 0.1Hz component to study other aspects of performance. For example, Weimann (1989) presented subjects with a novel, mathematical problem-solving task following a period of training. His results revealed a negative correlation between mental effort (as indexed by the 0.1Hz component) and time of solution for each problem. This finding suggested that subjects tended to reduce effort investment if the solution time for a problem was particularly protracted. A detailed investigation of effortful activity during applied performance was conducted by Tattersall & Hockey (1995). Eleven trainee flight engineers participated in a 3-hour working session in a simulated flight cockpit during this study. The various duties of the subjects were classified into three grades of task activity: (a) supervisory monitoring (e.g. low-level checking and routine maintanence), (b) fault rectification (e.g. detection and diagnosis of familiar fault states), and (c) problem-solving (e.g. detection and diagnosis of unfamiliar or abnormal fault states). This taxonomy was based on the tripartite classification produced by

(Rasmussen & Jensen, 1974) of skill-based, rule-based and knowledge-based levels of behaviour. For the purpose of the current discussion, we are concerned with the latter activity of problem-solving. These authors also employed the 0.1Hz component of HRV as an index of mental effort. Their results revealed higher levels of mental effort during problem-solving than either supervisory monitoring or fault rectification. This study corroborated the notion that higher levels of strategic mental effort may be necessitated during novel problem orientation than plan selection within a known task context.

The 0.1Hz component has also been used to study the influence of biological stressors such as stress and sleeplessness on performance in the field. For instance, 0.1Hz was used to quantify mental effort during a post-work battery of human performance tests (Mulders, Meijman, O'Hanlon, & Mulder, 1982). The authors found that effortful performance in the post-work phase (as indexed by 0.1Hz) was indicative of absenteeism and work-related stress. Similar studies have been conducted in recent times to study the relationship between work activity and mental effort for driving instructors (Meijman, 1995; Meijman, 1997).

The empirical chapters will use the 0.1Hz component of HRV in order to index psychophysiological effort. The following section will describe each individual study and its associated hypotheses.

3.4 Introduction to individual studies and associated hypotheses

Four experimental studies are described in the following four chapters. Each will focus on different aspects of the mental effort regulation model within the task context of sustained perceptual-motor behaviour.

The initial study described in Chapter 4 represents an investigation into individual differences, sleep deprivation and mental effort regulation. Based on performance during the control session, these subjects were divided into two groups of 'good' and 'bad' performers. The model would predict that good performance is achieved via mental effort investment coupled to an awareness of external performance. However, it may be difficult for those in the good performance group to sustain a successful policy of effort investment in the presence of sleep deprivation. This hypothesis is based on the assumption that sleep deprivation accelerates the magnitude of internal, energetical costs (e.g. reduced alertness), which requires compensatory effort at the expense of task-related effort.

Meanwhile, those in the poor performance group may invest effort at a significantly reduced level because their awareness of internal feedback is elevated relative to external feedback. It is anticipated that sleep deprivation will inflate this tendency and lead to a further reduction of task-related effort investment.

The first study hypothesised that sleep deprivation accelerated the evolution of energetical costs over a sustained period of task activity. This aspect of effort regulation was investigated further in the second study described in Chapter 5. This experiment was concerned with how the magnitude and category of biological stressor affected mental effort regulation. Therefore, subjects performed a sustained, perceptual-motor task in presence of different degrees of sleep deprivation (i.e. a night of 4 hours sleep compared to a full night without sleep). It is hypothesised that a full night without sleep would inflate energetical costs and lead to effort conservation relative to partial sleep deprivation, i.e. magnitude of stressor would be proportionate to the inflation of energetical costs. The model of effort regulation emphasises the significance of feedback and awareness of internal and external cues. This second study contrasted the influence of sleep deprivation with alcoholic intoxication in order to explore how biological stressors impinged on the process of feedback appraisal. It is postulated that feedback appraisal may be subject to perceptual bias due to this line of influence, which may limit the effectiveness of subsequent effort policy. In addition, the second study is concerned with the impairment of a real-world, perceptual-motor task – driving behaviour.

The issue of perceptual bias and feedback appraisal is explored further via the study described in Chapter 6. This experiment was also concerned with the role of effort regulation and driver impairment. It is postulated that the perception of external cues during driving may be distorted by the presence of stressors such as alcohol and fatigue. Furthermore, driving represents an example of skilled performance where the occurrence of errors may be low and/or individuals may have a tendency to appraise performance as superior to objective assessment. It was hypothesised that the provision of objective performance feedback would counteract these sources of perceptual bias. In addition, performance feedback should increase the salience of external cues and promote a policy of effort investment.

The final study (Chapter 7) will investigate the process of feedback appraisal during sustained performance under conditions of low and high task demand. It is hypothesised that demanding tasks emphasise external cues at the expense of internal feedback. However, if a task is undemanding, effort regulation tends to focus on internal feedback of energetical state. A second aspect of this study is concerned with the dual demands on effort regulation exerted by the task schedule (i.e. demand x duration of task activity). It is proposed that a demanding schedule will increase the antagonism between the need to invest effort and the desire to reduce effort. This aspect of effort regulation is explored via a manipulation of task schedule. In addition, this laboratory study extends the variety of subjective variables in order to present a more coherent index of energetical costs.

4 OPERATOR CAPABILITY, SLEEP DEPRIVATION AND SUSTAINED PERCEPTUAL-MOTOR PERFORMANCE[5]

4.1 Abstract

A study was conducted to investigate the influence of sleep deprivation on sustained perceptual-motor performance. Fifteen male subjects performed sixty minutes of tracking in combination with a secondary RT task under two conditions: (a) following a normal night of sleep, and (b) following a night without sleep. Data were collected with respect to primary task performance, psychophysiology and subjective indices. The subjects were divided into two groups of high- and low-competence performers with respect to tracking accuracy during the control condition (a). On this basis, two related hypotheses were tested: (i) are high-competence (HC) subjects characterised by higher investment of mental effort during sustained performance, and (ii) if so, does the greater expenditure of effort and associated costs put HC subjects at a disadvantage in the presence of an additional stressor such as sleep deprivation? Both hypotheses were contradicted by the results of the experiment. Low-competence (LC) subjects invested a higher level of mental effort regardless of experimental condition. This effect may have reflected a compensatory strategy in response to lower levels of electrocortical arousal as indexed by occipital alpha from the EEG record. By contrast, HC subjects were characterised by a higher level of performance efficiency. In addition, the pattern of costs associated with performance under both conditions revealed a number of significant differences between the two groups. The consequences of operator capability for effort policy are discussed, with a particular emphasis on the role of internal and external sources of feedback for the formation of effort policy.

4.2 Introduction

Few studies have addressed the interaction between individual differences and mental effort strategies. The paucity of this empirical data is unfortunate because effort policy incorporate important interactions between stable traits or habits and cognitive-energetic mechanisms. For example, individuals who inherently gravitate towards higher levels of goal aspiration may expend higher levels of mental effort relative to those with lower aspirations (Locke & Latham, 1990; Weiner, 1985).

[5] A partial report of this study was published as Fairclough, S. H., & Ward, N. J. (1996). A protocol for the assessment of subjective sleepiness. In H. F. Society (Ed.), *Human Factors and Ergonomics Society 40th Annual Meeting* (Vol. 2, pp. 1283). Philadephia: Human Factors Society..

According to Schonpflug (1986b), there are three major sources of individual differences which may influence mental effort policy: (a) personality traits, (b) acquired principles of mental effort regulation, and (c) the habitual focus of mental effort regulation.

The scale developed by Dornic, Ekehammar, & Laaksonen (1991) to measure tolerance for mental effort investment is the only example of a trait questionnaire (to the author's knowledge), where effort regulation is explicitly addressed as a stable source of individual differences. In most research on individual differences, effort regulation is only implicated in an indirect sense. For example, as a manifestation of underlying personality temperament (Strelau, 1985), or a factor within the interaction between personality traits and vigilance performance (Davies, Jones, & Taylor, 1984; Matthews, Davies, & Holley, 1993) or test anxiety (Eysenck, 1982; Eysenck, 1986). The link between personality traits and effort regulation was demonstrated by Schultz & Schonpflug (1982). These authors divided subjects into extroverts or introverts, and low and high anxiety subjects. During task performance, it was found that introverts and low anxiety subjects showed much slower levels of task disengagement (i.e. effort conservation) relative to extroverts and anxious subjects respectively.

Individuals may acquire specific principles of mental effort regulation based on previous experience. For example, extremes of the self-efficacy continuum are characterised by Bandura (1997) as contrasting perspectives on the nature of ability, i.e. efficacious individuals believe that ability is acquirable, whereas individuals with low self-efficacy assume that ability is inherent. The belief that ability may be acquired via persistence and experience may be a necessary precondition for mental effort investment under certain circumstances. By contrast, an assumption that all ability is inherent may provoke only a token investment of mental effort followed quickly by conservation if success is not apparent. Karasek (1979) hypothesised along similar lines, according to his model, highly competent individuals only tended to conserve effort under extreme demands (e.g. extreme fatigue or stress, repeated failure), whereas those less competent tended to adopt a strategy of effort conservation much earlier.

It may be hypothesised that all three groups (traits, principles, focus) may interact to influence mental effort policy. For example, a person predisposed towards test anxiety may habitually adopt a low-efficacy strategy based on indirect control, i.e. a sensitivity to internal feedback concerned with symptoms of sympathetic arousal. This cyclic chain of causality was inherent within the Self Regulatory Executive Function (S-REF) model proposed by Wells & Matthews (1994) (see Section 2.2.5 for description). Within the tripartite architecture of the S-REF model, the focus of effort regulation originates on the basis of lower-level feedback or from self-knowledge/beliefs at the upper level, which is translated into a coping response or principle of effort regulation at the middle level, which in turn, is based on stable traits inherent in upper-level self-knowledge. The difficulty associated with this conceptualisation is to pinpoint the proportionate influence of each of the three groups within this bi-directional cycle of causality.

The ability of an individual to sustain performance may be characterised according to their maximal capacity with respect to expenditure/replenishment ratio. This concept has been termed the 'coping capacity' of the individual, i.e. a set of resources which may be devoted to task performance (Schonpflug, 1983). This coping capacity of the individual is a similar concept to the zones of maximal adaptability prominent in the model of stress and performance proposed by Hancock & Warm (1989). This original conceptualisation was reformulated in Section 2.2.7 (Figure 10) as the capacity of the individual to respond to the demands of performance and to compensate for those energetical costs associated with performance.

The focus of the current experiment is on the principles and the focus of mental effort policy, particularly in relation to competence or proficiency. According to Schonpflug (1983), the competence of an individual refers to a repertory of mental and motor skills that are stable over time. The interaction between competence and mental effort expenditure is important, as the relationship between investment and performance may be mitigated by the competence of the individual. If a person is highly skilled, it may be predicted that less effort would be required to raise the level of performance effectiveness. Schonpflug (1983) also claimed that the total competence of the individual may define the maximal capacity of that individual. In other words, the limits of our coping capacity with respect to either performance or costs are determined by our level of competence.

In the same paper, Schonpflug proposed that the level of mental effort investment determined the proportion of total competence, which may be utilised in any given situation. Therefore, high competence should improve the efficiency and effectiveness of performance (Eysenck & Calvo, 1992) and this is the means by which competence may extend the limits of our adaptability to stress, i.e.

'stress' as used in a generic sense to describe the influence of task demand and environmental/biological stressors (Hancock & Warm, 1989).

The precursor to this study was an earlier investigation conducted by Schultz and Schonpflug (1979) and described in a later paper by the second author (Schonpflug, 1983)[6]. These authors had their subjects perform 'mental tasks' of differing levels of difficulty in the presence of different noise intensity. The subjects were consequently divided into three groups according to the frequency of errors (high, intermediate and low) and the data were re-analysed. This second analyses revealed no differences between the groups in terms of either task difficulty or noise level. However, the researchers discovered that the less competent group had the shortest response time even when making a correct response. In addition, these subjects were the least affected by either increased task demand or higher levels of noise. On this basis, it was concluded that subjects in the low competence group had adopted a conservation strategy, i.e. low output/low activity, which effectively insulated them from the influence of either noise or task demand. On the other hand, the high competence group sustained aspiration level and performance accuracy, but were forced to reduce their level of aspiration under the high noise condition. Therefore, the presence of noise had a substantive influence on this group because they were engaged in effortful activity.

The study performed by Schultz and Schonpflug illustrated how the level of competence of an individual may influence effort policy and therefore, exert a profound influence on performance in the presence of stressors. In this case, those who exerted mental effort found their coping capacity overextended in the presence of an environmental stressor. Paradoxically, those who were less competent did not suffer to the same degree.

The current study represents a partial replication of the Schultz and Schonpflug study. In this case, subjects performed a sustained, perceptual-motor task under two conditions, following a night of normal sleep or sleep deprivation. Therefore, subjects must regulate mental effort over a prolonged period of performance with or without the presence of a second stressor. In an early study of sleep deprivation, Wilkinson (1962) reported that subjects with the best performance following a period without sleep, exhibited a higher level of muscular tension. This study supported the view that subjects exerted higher levels of effort in order to compensate for the debilitating influence sleep deprivation on

[6] The original was only available in the original German at the time of writing.

performance. Therefore, in accordance with the logic of Schultz and Schonpflug's finding, if competence is associated with high effort expenditure per se, we may expect the presence of an additional stressor such as sleep deprivation to have a greater impact on the performance and effort policy of the most competent individuals.

4.3 Method

4.3.1 Experimental design.

The study was designed as a repeated measures design. The subjects performed an identical tracking task under two conditions: (a) following a night of normal sleep, and (b) following a night without sleep. In addition to the main effect of sleep deprivation, time-on-task (TOT) was measured by comparing data obtained during the initial fifteen-minute period of the task with data from the final fifteen minutes of the task. Two subject sessions were scheduled per day, one in the morning (10:00) and the other in the afternoon (14:00). All subjects participated at the same schedule time (either morning or afternoon) for each session. The order of presentation for (a) and (b) was counterbalanced across the subject group.

4.3.2 Apparatus and Experimental Task.

The perceptual-motor task utilised for this experiment was written in Think's C™ and ran on a Macintosh Quadra 700 with a 14" colour monitor. Subjects were required to track lateral target movements using the mouse. Specifically, the mouse controlled a white 'subject' cursor moving within a darker rectangle representing the software-controlled cursor. If the subject cursor made contact with the computer cursor, a tone sounded and the colour of the subject cursor turned from white to red. All subjects controlled the mouse with their right hand. A depiction of the computer screen is shown below in Figure 14. The speed of the tracking target was set to an intermediate level and its movements were determined by a random seeding process, i.e. that determined the frequency of lateral reverses performed by the computer-controlled tracking target. The movement of the computer-controlled cursor was stable across the whole task session.

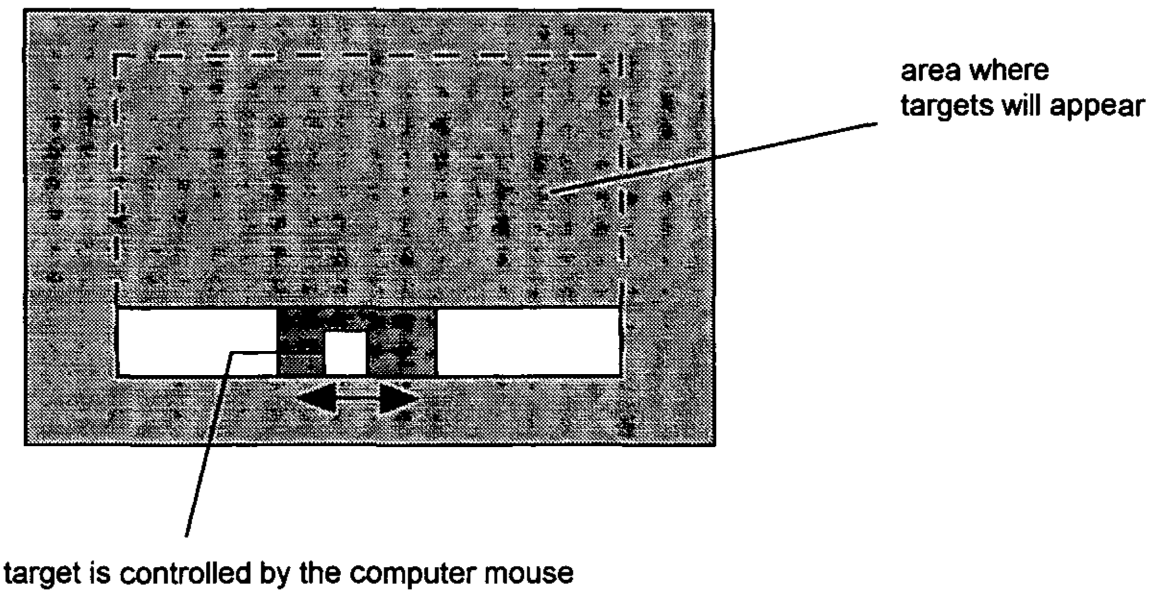


Figure 14. A representation of the screen image of the experimental task.

The subjects were also asked to perform a cognitive, vigilance task (Warm & Dember, 1998) in conjunction with tracking activity. The targets for this task appeared in the upper portion of the screen above the tracking task (Figure 14). The cognitive, vigilance task took the form of a letter identification activity. The subjects were presented with a Landolt C, which was shown in four different orientations (north, west, east, south). The subjects were instructed to recognise the easterly orientation as a target and all other orientations as non-targets (Figure 15). Subjects responded by pressing the mouse button and a response was only required in the presence of a target (Figure 15). In addition, each stimulus appeared onscreen for 600 msec, the event rate was 1.5 per minute and twenty percent of all RT stimuli were legitimate targets. The sampling rate used to collect data was 16000 Hz.

Figure 15. The stimuli used for the target detection task.

The subjects performed the perceptual-motor task continuously for sixty minutes. An eight-channel analogue-to-digital converter (MacLab™) was used to collect electroencephalograph (EEG) and electrocardiogram (ECG) data. This apparatus was connected to bioamplifiers and Chart™ software running on a Macintosh Powerbook™.

4.3.3 Subjects

Fifteen male subjects participated in the experimental study. The subject group were recruited from the local population and were paid for their participation. All subjects had 20/20 near visual acuity with the exception of two subjects, who had 20/40 acuity. The average age of the subject group was 32.4 years old (s.d. = 12.4, maximum = 59, minimum = 20). None of the subjects were shift workers nor on permanent medication. All subjects were asked to abstain from tea or coffee for 4 hours before the trial and only to consume a light meal before the test session.

4.3.4 Experimental Measures.

A number of data collection techniques were deployed during the study. These techniques have been divided into three distinct groups.

Primary task measures included data to index both tracking and RT performance. With respect to the former, these included: Root Mean Square error (RMS) and frequency of collisions, i.e. when the subject-controlled cursor made contact with the computer-controlled cursor. For the RT task, the level of target accuracy (% of hits, false alarms and misses) and mean reaction time (in ms) to targets and non-target were captured on the basis of raw data[7].

Psychophysiological measures constituted electrocortical arousal as indexed by the EEG record and heart rate variability from the ECG data. EEG data were collected at 100Hz from a single bipolar

[7] The level of RT accuracy was generally quite low. This finding is indicative of inadequate piloting of difficulty levels and instructions to subjects. Post-hoc feedback from subjects indicated that the event rate was too high and subjects elected to disregard the RT task in to sustain adequate tracking performance. It is difficult to assess the RT data as this task was largely ignored by subjects who had received no task prioritisation instructions. Therefore, these data were not analysed in the results section.

connection at $C_3 - O_2$ referred to A_1 at the left earlobe (International 10-20 System). The resulting data were subjected to a Fast Fourier Transform (FFT) analysis in order to calculate the level of power in three EEG bandwidths, i.e. theta (4-7Hz), alpha (8.5-12.5Hz), and beta (14-25Hz). These data were subjected to the bursts analysis described by Kecklund & Akerstedt (1993). Using this analysis, EEG power spectra were quantified for each consecutive 5 second period and baselined to pre-test performance. A burst of either theta, alpha or beta activity corresponded to the number of 5 second periods during every five minute period where energy spectra were equal to or greater than 125% of baseline.

The electrocardiogram trace (ECG) was recorded at 100Hz. The subsequent inter-beat interval (IBI) was subjected to power spectrum analysis in order to distinguish three bandwidths (upper, middle and lower). This analysis was performed using CARSPAN™ analysis software (Mulder & Schweizer, 1993). The mean power in the mid-frequency bandwidth (0.07 – 0.14Hz) was calculated to represent a psychophysiological index of mental effort (Aasman, Mulder, & Mulder, 1987; Mulder, 1979a; Mulder, 1979b; Vicente, Thornton, & Moray, 1987).

Subjective measures were administered during pre-test and post-test periods. These measures included the UWIST Mood Adjective Scale (Matthews, Jones, & Chamberlain, 1990), the raw version of the NASA-Task Load Index (Byers, Bittner, & Hill, 1989; Hart & Staveland, 1988), and a bipolar Effort scale (Zijlstra & van Doorn, 1985).

4.3.5 Experimental Protocol.

The subjects were asked to complete a number of consent forms in the week before participation and were instructed with respect to the sleep deprivation protocol. The subjects were informed that they would be expected to remain awake throughout the night prior to one of the two experimental sessions. The subjects verified their compliance with the sleep deprivation regime by telephoning an answering machine (which logged time and date) every hour between midnight and 8AM[8]. In addition, subjects signed a consent form in which they agreed not to drive a vehicle or engage in any activity that may be considered dangerous following a night of sleep deprivation (until they had obtained a period of sleep). Finally, the subjects were asked to specify if they wished to be brought to the Institute by taxi on the

[8] At 9AM, the subjects were contacted by the Institute every hour until their experimental schedule to confirm that they remained awake.

day of their sleep deprivation session, or if they wished to make their own transportation arrangement, i.e. to be brought and collected by a spouse, family member or friend.

On arrival at the Institute, the subjects were presented with written instructions, to describe the experimental task and the requirements of the study. The subjects perused this document whilst EEG and ECG electrodes were attached. Subjects were then seated in a separate room and observed by the experimenter via a one-way mirror. An intercom was used in order to communicate with the subject. The subjects were allowed to practice the experimental task for fifteen minutes. During the control condition, this period was used to collect baseline psychophysiological data. When the training session had concluded, the subjects completed a pre-test index of subjective measures. Once the questionnaires had been completed, an opportunity was provided to resolve any queries or uncertainties on the part of the subjects.

The subjects performed the tracking task continuously for sixty minutes in both conditions. Once the task was terminated, the subjects were required to complete a post-test index of subjective questionnaires. The subjects' electrodes were removed following this point and the subjects were both debriefed and paid or scheduled for their next session. There was a period of seven days between each subject session.

4.4 Results

4.4.1 Classification of high- and low-competence subjects.

The subject group was divided into two groups of six high-competence and low-competence individuals. This distinction was based on tracking performance during the control session (i.e. following a night of normal sleep), specifically on the number of collisions made by each subject. High-competence (HC) subjects were classified as those subjects who made 60 collisions or less during the control condition. Low-competence (LC)subjects were identified as those individuals who made 120 collisions or more. There were six subjects in each group. Three subjects who fell into the >60 and <100 collisions zone were omitted from all future analysis. A frequency distribution chart for these data is shown below in Figure 16.

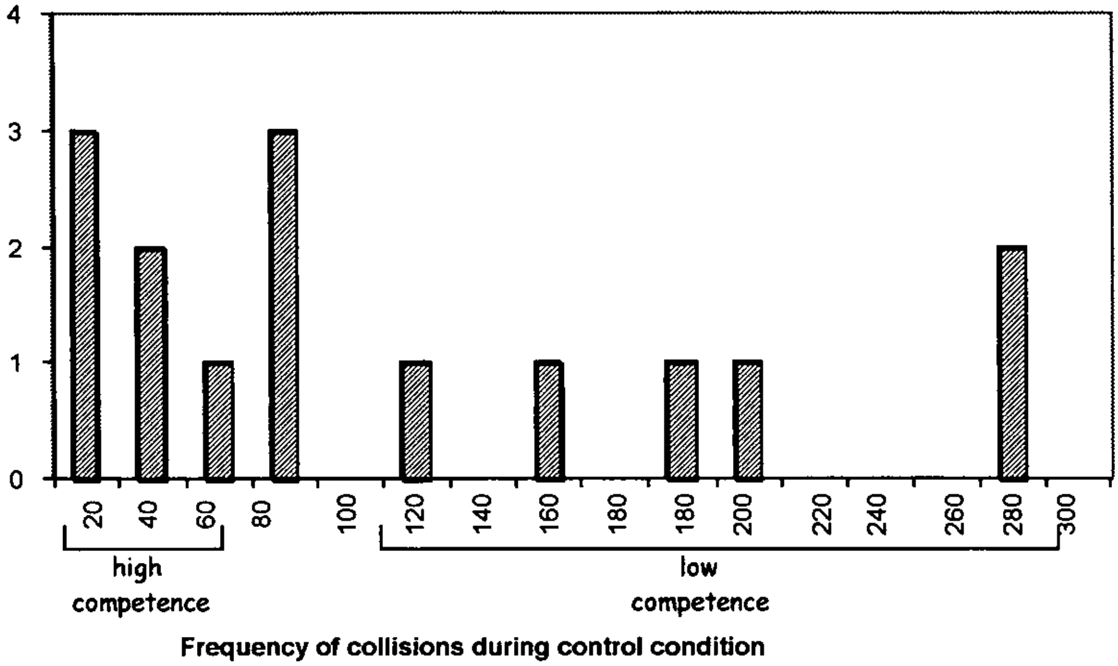


Figure 16. Frequency distribution for total number of collisions during the control conditions (N=15).

This classification of the subjects was checked against experimental scheduling information to check for possible confounding due to the time-of-day variable, i.e. given that half the subjects were tested in the morning and half during the afternoon. No evidence was found that time-of-day exerted an influence over the membership of either subject group, i.e. four of the HC group were tested during the afternoon, whereas LC subjects were equally split between morning and afternoon sessions. In addition, a non-parametric Mann-Whitney test was performed to test for age differences between the two groups. No significant age differences were found between the two groups.

4.4.2 Primary task performance

These data were averaged during the initial and final fifteen-minute period of each experimental session, i.e. to investigate time-on-task effects (TOT). A 2 x 2 x 2 ANOVA (GROUP x CONDITION x TOT) was performed to confirm the group classification and to investigate the influence of other main effects for the frequency of collisions. As anticipated, these analyses revealed a significant main effect due to GROUP [$F(1,10)=7.0$, $p < 0.05$] in the expected direction. In addition, both CONDITION [$F(1,10)=13.9$, $p < 0.01$] and TOT [$F(1,10)= 9.9$, $p < 0.05$] were significant main effects, i.e. collision frequency significantly increased in the presence of sleep deprivation and due to TOT.

An identical analysis was applied to the RMS error data in order to consider tracking performance not exclusively associated with error. This ANOVA revealed identical significant main effects due to GROUP [$F(1,10)=9.9$, $p < 0.05$], CONDITION [$F(1,10)= 20.4$, $p < 0.01$], and TOT [$F(1,10)=26.9$, $p < 0.01$] as the previous analysis. In addition, a significant 2 x 2 x 2 interaction was present [$F(1,10)=4.9$, $p = 0.05$]. Post-hoc testing (via Tukey HSD) indicated that RMS error was not significantly different during the initial period of the control and sleep deprivation sessions for highly competent subjects.

4.4.3 Psychophysiology

A series of 2 x 2 x 2 ANOVAs were performed to explore the influence of the three main effects on the frequency of theta, alpha and beta bursts. No significant effects were observed in the analysis of theta bursts. The analysis of alpha bursts revealed a significant main effect due to GROUP [$F(1,10)=5.1$, $p < 0.05$], which indicated that the frequency of alpha bursts was significantly higher for the low competency subjects. This effect is illustrated in Figure 17. There was also a significant main effect for TOT [$F(1,10)=8.1$, $p < 0.05$], i.e. the frequency of alpha bursts tended to increase between the initial and final period of task performance. The analysis of beta bursts revealed only a significant main effect for TOT [$F(1,10)=14.2$, $p < 0.01$], i.e. the frequency of beta bursts declined between the initial and final period of task performance.

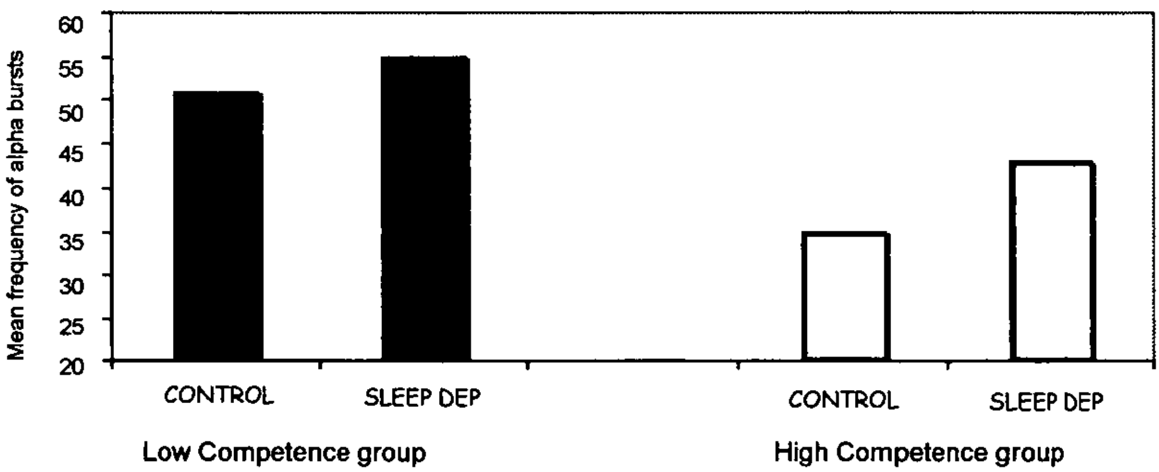


Figure 17. Mean frequency of alpha bursts (per 5 minute period) between high and LC subjects in both conditions (N=12).

The mean power in the middle bandwidth of heart rate variability was subjected to log transformation and baselined to the training session of the control condition. A 2 x 2 x 2 ANOVA revealed a main effect of marginal significance for GROUP [$F(1,10)=4.3$, $p = 0.08$]. This effect indicated that the LC subjects exerted a higher level of psychophysiological effort relative to the HC group, regardless of experimental condition. In addition, there was a significant main effect due to TOT [$F(1,10)= 43.2$, $p < 0.01$], i.e. the level of psychophysiological effort was significantly lower during the final fifteen minutes of performance.

4.4.4 Subjective data

The data from the subjective index of questionnaires were subjected to 2 x 2 x 2 ANOVA analyses (GROUP x CONDITION x TOT). In this particular case, TOT was measured in terms of pre- and post-test scores on all subjective scales.

The analysis of the bipolar Effort scale revealed a significant main effect due to TOT [$F(1,10)= 11.9$, $p < 0.01$], i.e. subjective estimates of mental effort were higher following task activity ($M = 62.9$) than during the pre-test period ($M = 46.3$). In addition, there was a significant 2 x 2 interaction between GROUP and CONDITION [$F(1,10)= 5.2$, $p < 0.01$]. Post-hoc testing revealed that HC subjects perceived a higher level of mental effort during the sleep deprivation condition, whereas the LC group perceived no distinction between the experimental conditions with respect to subjective effort. This interaction is illustrated below in Figure 18.

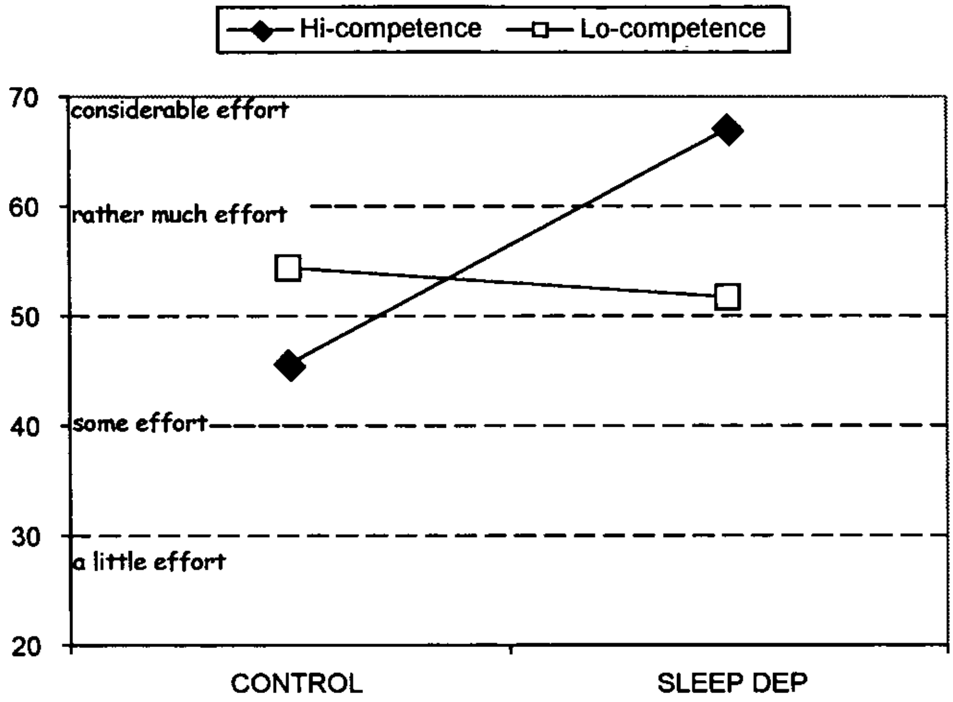


Figure 18. Mean ratings of subjective effort from the subjective effort scale (Zijlstra & van Doorn, 1985) for each experimental condition for both groups of subjects (N=12). Note: written labels included from the scale.

The analysis of the raw NASA-TLX data revealed a number of significant effects. The results of the TLX analyses are shown below in Table 5. It was apparent that subjects perceived an increase of mental workload between the pre- and post-test administrations. In specific terms, subjective estimates of mental demand, physical demand and frustration all increased over TOT.

Variable	GROUP A	CONDITION B	TOT C	INTERACTIONS
Mental demand			F(1,10)=18.6 p < 0.01	A x B x C F(1,10) = 7.5 p < 0.05
Physical demand			F(1,10)=21.2 p < 0.01	
Temporal demand				A x B x C F(1,10) = 14.3 p < 0.01
Performance				A x B F(1,10) = 4.4 p < 0.05
Effort				
Frustration	F(1,10) =19.3, p < 0.01		F(1,10) = 7.9 p < 0.05	
Subjective mental workload			F(1,10) = 18.3 p < 0.05	

Table 5. Results of Raw Task Load Index (RTLX) analyses (N=12).

The only main effect due to the GROUP variables observed during the RTLX analysis was the finding that LC subjects experienced a higher level of frustration, regardless of experimental condition, compared to the HC group.

A number of interaction effects were found during the analyses of the RTLX as shown in Table 5. The post-hoc testing of the mental demand data revealed that pre-test ratings of the LC subjects were significantly lower than their HC counterparts, but only during the sleep deprivation condition.
A similar effect was observed with respect to temporal demands. This interaction is shown in Figure 19. Post-hoc testing revealed that pre-test ratings of temporal demands were lower for LC subjects relative to the HC group during sleep deprivation. In addition, LC subjects perceived temporal demands to increase during the sleep deprivation session, whereas HC subjects did not. It was also apparent that post-task ratings of temporal demands increased for the high competent subjects ($p < 0.05$).

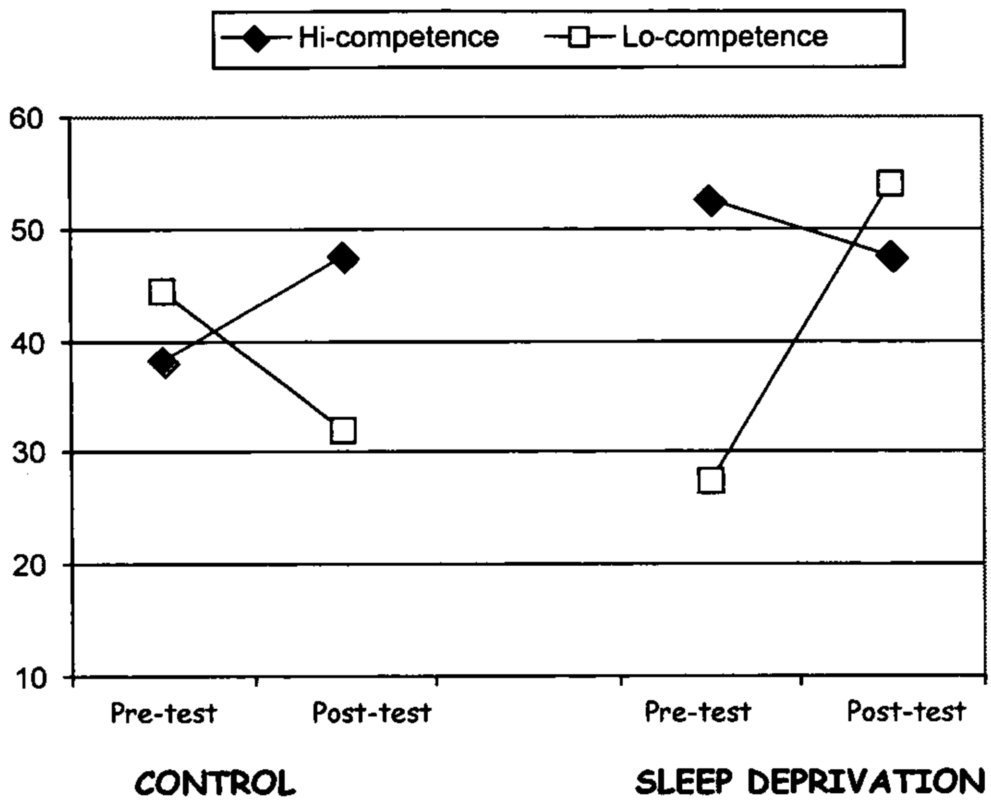


Figure 19. Subjective ratings of temporal demands from the RTLX for both groups across both conditions (N=12).

The operator performance sub-scale of the RTLX assessed subjects' perceptions of their own level of performance. The significant interaction between GROUP x CONDITION (Table 5) was subjected to post-hoc analysis, which revealed that HC subjects perceived their performance to be significantly superior in the control condition but worse following a night without sleep. However, the LC group did not perceive any similar decline due to sleep deprivation. This interaction is shown in Figure 20.

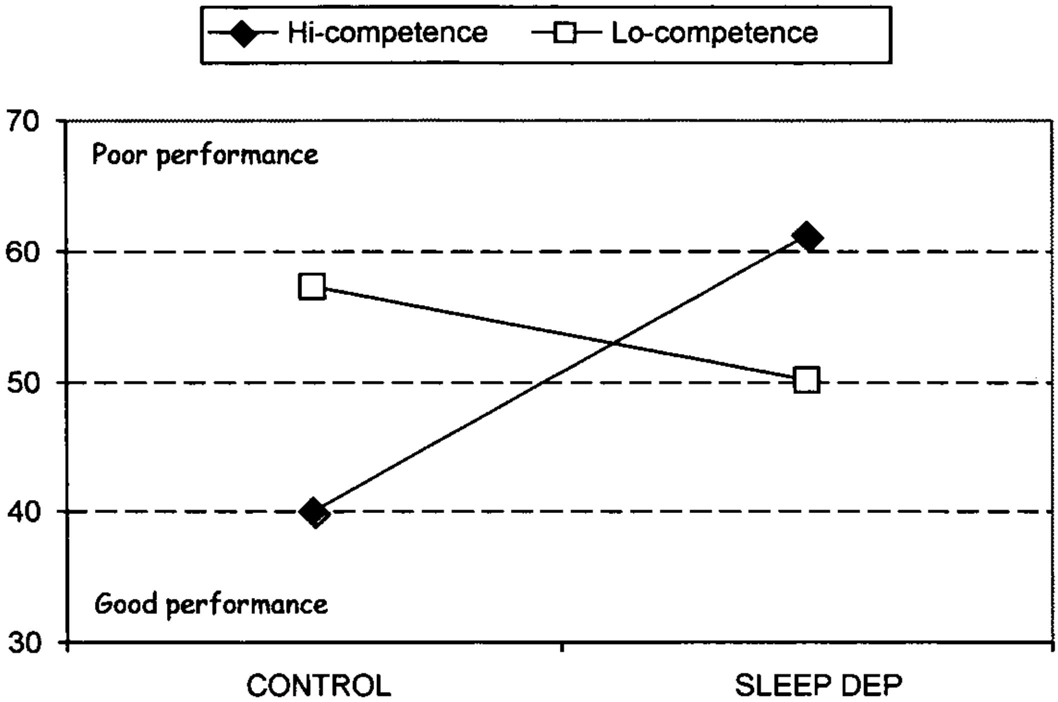


Figure 20. Subjective ratings of operator performance from the RTLX for both groups across both conditions (N=12).

The UWIST mood scale was analysed in terms of three principle components: energetic arousal (EA), tense arousal (TA), and hedonic tone (HT). The analysis of EA revealed only a significant interaction of GROUP x CONDITION. Post-hoc testing indicated that HC subjects experienced a decline of EA due to sleep deprivation whereas the LC group did not. The analysis of TA revealed a significant main effect of CONDITION [$\underline{F}(1,10) = 5.9$, $\underline{p} < 0.01$], i.e. tense arousal was significantly higher in the sleep deprivation condition, but no other significant effects. With respect to the final mood component, it was found that HT was significantly reduced due to TOT, i.e. all subjects experienced higher negative affect with increasing TOT.

4.4.5 Performance Efficiency Index.

The efficiency of performance was operationalised by representing the relationship between psychophysiological effort and primary task performance. This representation is based on the analysis described by (Meijman, 1995).

The 0.1Hz data were expressed as a percentage change from initial data collected during the training session of the control condition, i.e. increasing positive percentage values being indicative of increased mental effort (Meijman, 1995). The quality of tracking performance was represented by the number of collision for each group. Both groups of data were quantified at 5-minute intervals (excepting the initial 5-minute period) and are plotted for both groups across both conditions in Figure 21.

The efficiency analysis in Figure 21 illustrates how sleep deprivation made a more substantive impact on tracking performance rather than the level of mental effort. It is also apparent that high competent subjects were characterised by a higher level of efficiency (i.e. lower mental effort, superior performance) regardless of the experimental condition. It should be noted that sleep-deprived, HC subjects generally achieved a higher level of efficiency than fully rested, LC subjects.

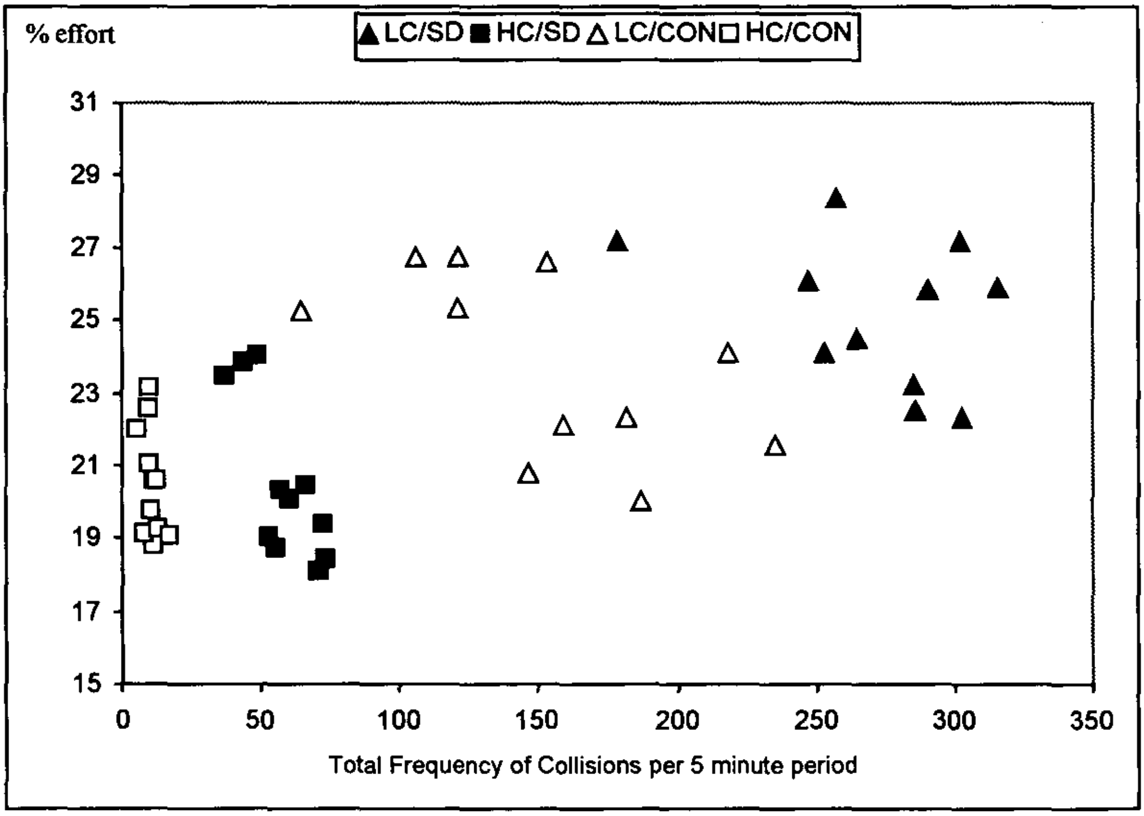


Figure 21. Performance efficiency as represented by the relationship between tracking performance (frequency of collisions) and psychophysiological mental effort for both groups across both conditions (N=12). Note: high x-values are synonymous with poor performance whereas high y-values are associated with an increase of mental effort.

4.5 Discussion

The data analysis revealed that the lower level of tracking performance exhibited by LC subjects was accompanied by a decline in electrocortical arousal (Figure 17) and an increase in psychophysiological effort relative to the HC group. These findings contradict one of the main hypotheses of the study, i.e. that HC subjects attain superior performance via higher effort investment.

The subjective data indicated that LC subjects were less affected by the sleep deprivation manipulation than the HC group. For instance, LC subjects did not mobilise subjective effort in response to sleep deprivation (Figure 18). In addition, LC subjects perceived neither a decline of energetic arousal nor a fall in the quality of performance (Figure 20) when deprived of sleep. This insensitivity to the sleep deprivation manipulation contrasted to the perceptions of HC subjects, who generally assessed a decline of energy and performance and sought to invest subjective effort as a means of compensation. Therefore, the magnitude of energetical costs associated with performance showed a relative increase due to sleep deprivation, but only for HC subjects.

There is a caveat to this statement concerning the main effect due to GROUP associated with perceived frustration (Table 5). It was apparent that LC subjects perceived their level of frustration to be higher than those in HC group regardless of the experimental condition. This cost was not accompanied by any other affective change (from the UWIST Mood Scale), however, it does provide circumstantial evidence that either: (a) LC subjects were aware of their frequent error count and were frustrated by their inability or unwillingness to improve performance, or (b) LC subjects were less comfortable than HC subjects within the test scenario per se and experienced frustration due to test anxiety (Eysenck & Calvo, 1992). With respect to hypothesis (a), it was apparent that LC subjects assessed their performance to be inferior to HC subjects, but only during the control condition (Figure 20).

The analyses of performance, psychophysiology and subjective measures indicated that the poor performance of LC subjects was associated with a higher rate of effort expenditure. However, this additional effort investment did not improve performance effectiveness (Figure 21). Therefore, it is hypothesised that LC subjects exhibited a higher rate of effort expenditure in order to compensate for reduced electrocortical arousal (Figure 17). Psychophysiological evidence from the EEG supported this conclusion, the frequency of alpha bursts was significantly higher for LC subjects regardless of sleep deprivation. It may be postulated that an increased susceptibility to boredom or daytime sleepiness may represent candidate trait variables responsible for the poor performance exhibited by LC subjects. A number of trait questionnaires have been produced such as the Boredom Proneness scale (Farmer & Sundberg, 1986) and the Epworth scale to quantify susceptibility to daytime sleepiness (Johns, 1992). In addition, the former scale has been linked to individual differences with respect to performance in the vigilance domain (Sawin & Scerbo, 1995). It is important to note that both increased psychophysiological effort and reduced electrocortical arousal characterised the LC subjects in both experimental conditions. This factor may be presented as evidence that a stable trait variable was responsible for differences between LC and HC subjects.

In the Schultz and Schonpflug study (Schonpflug, 1983), LC subjects were characterised by an acquired principle of effort conservation. When these subjects encountered resistance from a stressor, they systematically reduced levels of aspiration and activity with a consequent decline of performance. The LC subjects in the current study were characterised by poor performance in combination with relatively higher levels of psychophysiological mental effort, alpha activity and subjective frustration. These multidimensional indicators indicate that LC subjects in the current study were characterised by a highly inefficient strategy of mental effort investment, where maximum effort is paired with performance of minimal effectiveness.

The effort policy enacted by the LC group may be described as inefficient investment in both experimental conditions. For example, the efficiency index shown in Figure 21 illustrates how sleep-deprived HC subjects performed with greater efficiency than fully rested LC subjects. These findings indicate that the relationship between performance capacity and effort investment was different for HC and LC subjects. Given that performance capacity may be described in terms of trait and state variables, there are two possible explanations for this distinction: (a) that the HC group had an advantage with respect to trait characteristics such as superior sensory capabilities or perceptual-motor ability, (b) that the LC group were at a disadvantage due to transient state variables, or (c) a combination of both trait and state factors. The hypothesis (a) is almost impossible to investigate on a post-hoc basis and seems unlikely (given subjects were tested for visual acuity and no age differences were found between the two groups), but cannot be ruled out completely. However, there was evidence to support an influence of state variables as described in hypothesis (b). Based on the analysis of psychophysiology, it was apparent that LC subjects must invest effort in order to compensate for higher levels of alpha in the EEG record, i.e. which has been linked with a general degradation of sustained performance (Gale, 1977; O'Hanlon & Beatty, 1977; Thackray, Bailey, & Touchstone, 1977), however this link may not be consistent (Parasuraman, 1983).

This investment of compensatory mental effort may effectively limit the maximum performance capacity of those LC individuals. Given that maximum capacity is influenced by the level of task-related effort (Schonpflug, 1983), the diversion of mental effort to compensate for high alpha activity may ultimately represent a limitation on effort investment into task-related activity. On the other hand, the HC subjects do not have to compensate for this category of energetical cost to the same extent and a higher proportion of the effort invested by HC subjects is devoted to task activities, resulting in superior levels of performance efficiency. A diagram to represent this relationship between effort and capacity is shown below in Figure 22, based on the analysis of Schonpflug (1983). Therefore, efficient HC subjects must only invest 50% of mental effort to achieve maximum capacity. Whereas the LC subjects require 100% of the available mental effort to achieve a lower level of performance (relative to HC subjects), i.e. to perform the task and to compensate for increased alpha in the EEG.

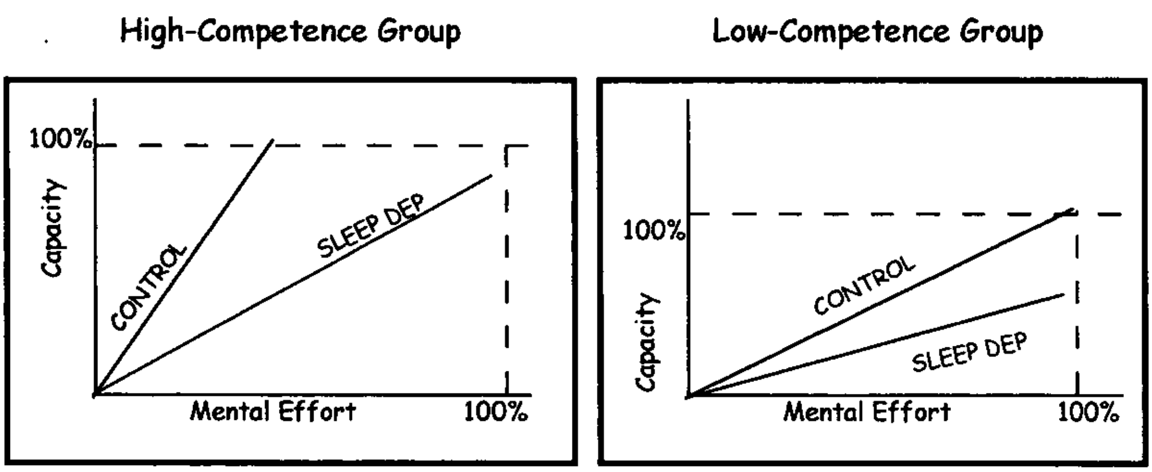


Figure 22. Hypothetical relationship between capacity and mental effort for both groups of subjects.

It was noted that the HC subjects showed greater sensitivity to the subjective costs associated with sleep deprivation than the LC group, e.g. reduced energy, poorer performance, higher level of effort investment. It was also found that LC subjects perceived mental and temporal task demands to reduce during the pre-test period of the sleep deprivation session (Figure 19). This perceived decline of task demands might indicate a reduced sensitivity to the increase of workload induced by sleep deprivation.

These interaction effects may be used as evidence that: (a) LC subjects were unaware of the influence of sleep deprivation on energy and performance, therefore, made no attempt to increase effort, or (b) that awareness of low energy and poor performance peaked during the control condition for the LC subjects, and these individuals were subsequently insensitive to the supplemental demands of sleep deprivation. This latter hypothesis is illustrated in Figure 22, i.e. the difference between both conditions is minimal for the LC subjects. This increased separation for the HC group is intended to represent the higher adaptability of these subjects, i.e. they have the option of several degrees of effort investment above the control condition. On the other hand, the only effort policy available to the LC subjects is one of maximum effort investment, regardless of the presence of sleep deprivation. The lack of adaptability or control over effort regulation for those LC subjects may account for a reduced sensitivity to those additional costs introduced by sleep deprivation. This analysis emphasises the link between performance efficiency and adaptability of effort policy. A similar analysis was performed by Hancock (1986) to illustrate the interaction between competence and resistance to heat stress during performance, i.e. competent subjects were more capable of resisting the debilitating influence of heat stress on performance.

It is postulated that a fundamental link exists between the two hypotheses (a) and (b) - the connection between awareness (of feedback) and adaptability of effort policy. According to the model of effort regulation in Chapter 2, the adoption of a mental effort principle is based on an assessment of efficiency, originating from external (task performance) and internal (self) sources of feedback. The mechanisms underlying this model are based on control theory, therefore, the individual must be aware of feedback in order to exert control. The same logic would predict that an awareness of feedback (with respect to either external or internal sources of information) is a necessary precondition of adaptive effort regulation. This chain of causality indicates how reduced awareness of feedback may undermine adaptive effort regulation, with respect to either hypothesis (a) or (b) in the previous paragraph.

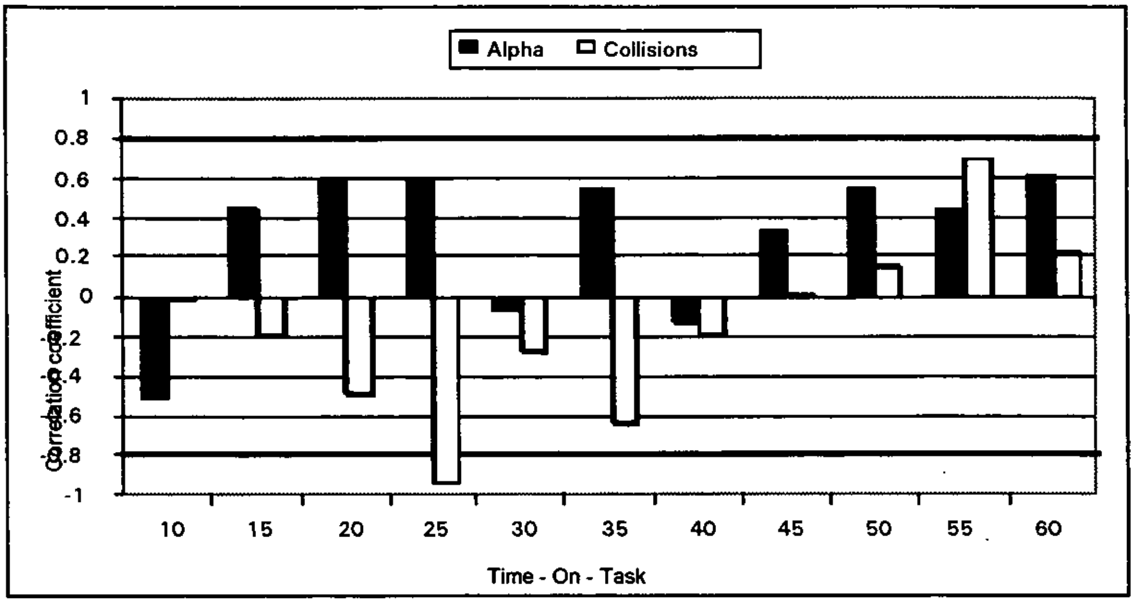
In order to investigate this issue in detail, a supplementary analysis was performed to explore how psychophysiological effort was regulated by each of the two subject groups. Two particular influences on effort regulation were included in this correlational analysis:

- alpha bursts (internal feedback)
- the frequency of collisions (external feedback)

It was hypothesised that increased alpha activity might prompt effort investment in order to compensate for energetical costs based on internal feedback. With respect to external feedback, a collision between subject- and computer-controlled cursor was associated with overt visual and auditory feedback of performance effectiveness. A series of correlation coefficients were quantified to describe the pattern of association between data pairs based on consecutive five-minute windows. These analyses were based on baselined values of psychophysiology and performance. The latter were baselined to the initial five-minute period of the control condition, hence only eleven of the twelve possible five-minute periods per session were included in the analyses.

The analysis of correlated data is illustrated in Figure 23 for the control condition. The pattern of data for the HC subjects (Figure 23a) indicates that psychophysiological effort generally had a positive correlation with the frequency of alpha bursts. Although these coefficients failed to reach significance, the pattern is consistent throughout the session. There is some indication that increased effort was associated with a reduction of collisions between 20 and 35 minutes, but this trend was sporadic. It is difficult to detect any consistent pattern of correlation for the LC subjects in Figure 23b. There are some indicators of negative correlation between effort and collisions/alpha bursts, but these trends are inconsistent and sporadic. The surprising aspect of these supplementary analyses was the association between alpha activity and effort for the HC subjects. Based on previous data analyses, it would be expected that LC subjects would exhibit this trend, as this group experienced the higher level of alpha activity. In addition, it was anticipated that compensation for increased alpha would be more prominent during the sleep deprivation condition.

(a)



(b)

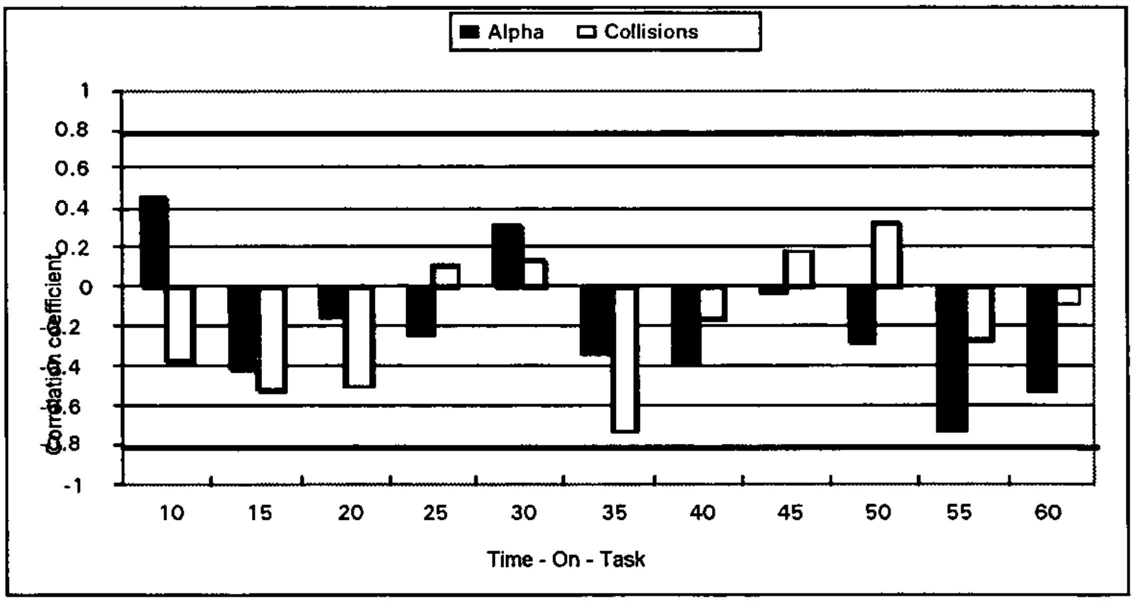
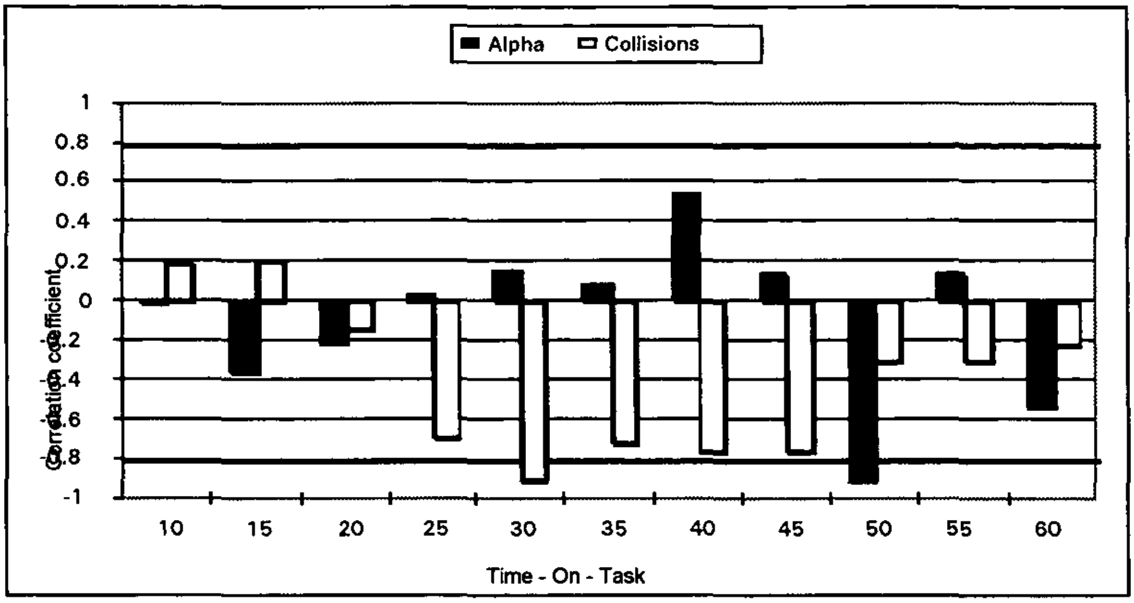


Figure 23. Correlation coefficients for psychophysiological effort and EEG/performance for (a) High-Competent subjects and (b) Low-Competent subjects during the Control condition (N=12). NB: R values above 0.8 are significant at $\underline{p} < 0.05$.

(a)



(b)

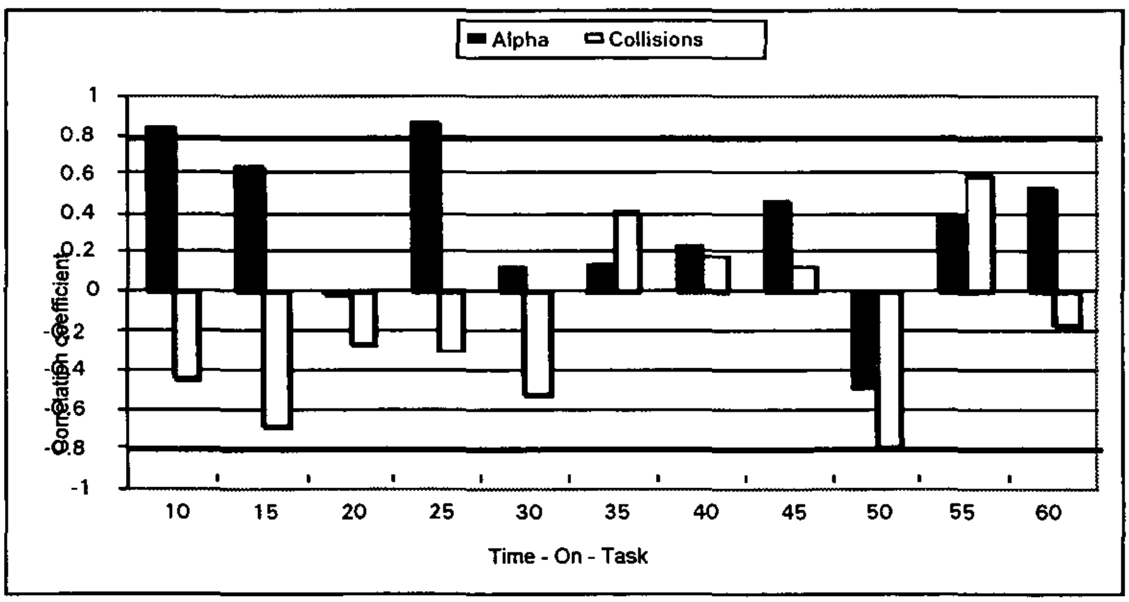


Figure 24. Correlation coefficients for psychophysiological effort and EEG/performance for (a) High-Competent subjects and (b) Low-Competent subjects during the Sleep Deprivation condition (N=12). NB: R values above 0.8 are significant at $\underline{p} < 0.05$.

The same analysis was conducted on data from the sleep deprivation condition (Figure 24). In this particular condition, the HC subjects appear to use external feedback from collisions as their primary spur for effort investment (Figure 24a), i.e. between 25 and 45 minutes. The pattern of negative correlation indicates that those HC subjects with the highest effort investment produced the lowest number of collisions or vice versa. The pattern of correlational data for the LC subjects shown in Figure 24b is not indicative of any consistent pattern. There is some evidence that effort investment increased in line with frequent alpha bursts between 10 and 25 minutes, and consequently reduced the number of collisions – but this trend was not sustained. As in the previous analysis, the pattern of correlational data was not in the expected direction. It was anticipated that effort would be invested in order to compensate for increased alpha activity due to the sleep deprivation. However, the only consistent pattern indicated a relationship between effort investment and external feedback from tracking performance (Figure 24a).

The correlational analysis should be treated with caution, the subject numbers in each group were very low and therefore, any pattern of association is highly unstable. In addition, it has been argued that effort regulation is based on the perception of internal and external feedback (Section 2.2.3), whereas this analysis was based on 'raw' sources of energetic and performance data. It would be unjustifiable to suggest that actual increases in alpha activity or collisions are necessarily perceived by the individual, e.g. the LC subjects do not perceive performance to decline or alpha to rise during the sleep deprivation condition, however, objective data sources indicated otherwise. Both caveats should be borne in mind during the following discussion of the correlation analyses.

The pattern of correlation coefficients for the HC group presented in Figure 23a and Figure 24a provided some evidence that the effort regulation was performed in line with external and internal sources of feedback. However, this pattern was surprising in the sense that effort was regulated in line with alpha activity during the control condition, and in response to collisions during sleep deprivation – the opposite trend would have been predicted. In addition, it was apparent that one source of feedback was used almost exclusively in each experimental condition, i.e. subjects did not seem to switch between internal and external sources during different phases of the experimental session. On this basis, it is hypothesised that either external or internal sources of feedback may be prioritised for effort regulation. It is logical to suggest that prioritisation may be determined by the source of feedback most closely associated with the most salient form of stress (Hancock & Warm, 1989). During the control condition, the HC subjects may perceive the most significant threat for sustained performance to originate from electrocortical arousal, i.e. that the ability to sustain alertness was threatened to a greater extent that the ability to sustain performance during the control condition. However, the introduction of sleep deprivation failed to reinforce this trend. It was apparent from the subjective data analyses, that sleep deprivation increased mental workload as well as reducing HC subjects' levels of energy. Therefore, it is reasonable to assume that sleep deprivation threatened the performance capability of the HC subjects, which led to a shift of effort regulation to emphasise external feedback from task performance.

The analysis of the LC subjects correlation data shown in Figure 23a and Figure 24a did not exhibit any consistent pattern. These data appear to contradict the earlier hypothesis that LC subjects were less efficient due to their need to compensate for increased alpha activity. No evidence for this compensatory effort was found during the correlational analysis. Therefore, it may be suggested that LC subjects increased effort investment without any clear focus for regulation. This hypothesis is suggested by their low awareness of either external or internal sources of feedback, both in the correlational analyses and with respect to subjective data. Therefore, the LC subjects exhibit a higher level of effort expenditure that fails either to improve performance or to alleviate energetic costs.

One major drawback with this study was the small size of the original subject sample. Each subject group only constituted six individuals in total, which is sufficient to threaten the integrity of the correlational analyses. In addition, the distinction between the two subject groups appeared to be uneven. The LC subject group encompassed a huge range of performance variability relative to the HC subjects (Figure 16). A larger subject sample would have permitted a more cogent means of partition for subject discrimination. The use of multidimensional measures to characterise effort policy represents an improvement on the method used by Schultz and Schonpflug (Schonpflug, 1983), where strategy was assumed based on primary task measures alone. However, the inclusion of real-time measures to capture subjective perceptions of performance, aspiration and internal costs would have provided the opportunity to investigate the level of association between perceived feedback and actual feedback.

4.6 Conclusions

The experiment indicated that LC subjects invested a higher level of mental effort and experienced a significant decline of electrocortical arousal relative to the HC group. This trend was apparent regardless of the experimental condition and no evidence was found that effort was invested to compensate for increased alpha in the EEG record. Therefore, it may be concluded that LC subjects were characterised by a stable trait variable, which was not responsive to state changes. These conclusions suggest that the poor performance of LC subjects resulted from either: (a) a source of degraded perceptual-motor ability that remained impervious to effort investment, or (b) an insensitivity to internal and external sources of feedback, which resulted in an inefficient and ineffectiveness effort policy.

The HC subjects exhibited superior performance effectiveness and efficiency in conjunction with lower alpha activity in both experimental conditions. These subjects appeared to have two primary advantages over the LC group:

- The HC subjects were capable of a higher level of performance efficiency. This superior level of efficiency was responsible for maximising coping capacity. In practical terms, reduced efficiency may aid the individual by expanding his or her adaptability to sources of stress. Therefore, the HC subjects were able to sustain relatively low levels of error in the presence of sleep deprivation without notable effort investment (Figure 21). This conclusion contradicts the hypotheses that HC subjects achieve proficiency via an increase of mental effort and suffer increased costs in the presence of an additional stressor. According to the current study, proficiency is simultaneously a cause and consequence of increased performance efficiency.

- The HC subjects exhibited a rationale allocation of mental effort expenditure based on an awareness of both external and internal feedback. During the control condition, HC perceived the greater threat to performance to originate from increased alpha activity and responded accordingly. When the HC subjects were sleep-deprived, collisions were interpreted as the principle symptom of increased drowsiness and effort was invested in line with external feedback. In addition, the HC subjects exhibited increased sensitivity to the influence of sleep deprivation on workload and energetical costs, relative to the LC group. It is hypothesised that awareness of feedback played an important role to ensure maximum levels of performance efficiency.

It is concluded that competence plays an important role with respect to capacity to sustain performance in the presence of a stressor. However, it is difficult to pinpoint whether competence results from trait factors, habitual effort strategies or the awareness of feedback. For example, does competence arise from higher achievement motivation, increased self-efficacy or a superior awareness of internal and external feedback? As described by Wells & Matthews (1994), these factors are cyclic and causality is difficult to establish.

The model of effort regulation postulates that awareness of internal and external feedback is central to an appraisal of efficiency, which drives future effort policy. This study has provided some supporting evidence for this model, with respect to the significance of external and internal feedback. It was suggested by the correlational analyses that HC subjects focused on external feedback at the expense of internal feedback and vice versa during different conditions. This hypothesis suggests that internal and external sources of feedback may be prioritised in order to permit a consistent formulation of effort policy.

5 IMPAIRMENT OF DRIVING PERFORMANCE DUE TO SLEEP DEPRIVATION OR ALCOHOL: A COMPARATIVE STUDY[9].

5.1 *Abstract*

A study was conducted to assess the relative impact of partial sleep deprivation (restriction to 4 hours sleep prior to testing) and full sleep deprivation (no sleep on the night prior to testing) on simulated driving, compared with an alcohol treatment (mean % Blood Alcohol Content = 0.07). It was predicated that both sleep deprivation and alcohol would influence internal and external sources of feedback and influence subsequent effort policy. The level of task demand was varied within the two-hour journey. Sixty-four male subjects participated in the study. Three categories of data were collected from the primary driving task, psychophysiology and subjective self-assessment. The results revealed that a full night of sleep deprivation produced the greatest magnitude of performance impairment, followed by the alcohol manipulation, whereas the performance of the partial sleep deprivation group was equivalent to control subjects. In addition, the partial sleep deprivation group exerted the highest level of psychophysiological effort. These results are discussed with reference to performance impairment, self-appraisal and the role of effort policy.

5.2 *Introduction*

It is significant that the dangers associated with driver fatigue have not found equivalent expression in the public arena as related concerns, such as drunk driving. The reasons why fatigue has traditionally played a secondary role in road transport safety research and policy were outlined by Brown (1994), namely: (a) that evidence of a causal influence of fatigue in accidents is often circumstantial, (b) a lack of political will to research the limitation of the hours of work for professional drivers, but underpinning both is the fact that, (c) fatigue is a complex and ambiguous concept with no standard measurement index. The absence of definition lies at the crux of the fatigue problem, in terms of both research and policy.

---

[9] This study has been published as Fairclough, S. H., & Graham, R. (1999). Impairment of driving performance caused by sleep deprivation or alcohol: A comparative study. *Human Factors*, 41(1), 118-128. v41 (1), pp. 118-128.

In the case of impairment due to alcohol, the Blood Alcohol Content (BAC) measure constitutes a broadly linear scale (based on physiological data) with a demonstrated correspondence to road accident involvement (Walls & Brownlie, 1985). Despite a long and varied research history (Bartlett, 1953; Bartley, 1965; Broadbent, 1979; Craig & Cooper, 1992; Crawford, 1961; Muscio, 1921; Thorndike, 1900), no equivalent index of fatigue has been forthcoming. This failure may be simply a conceptual limitation. Fatigue, like any affective construct such as happiness or anger, is a state spontaneously perceived by the individual, in the absence of any explicit verbal or numeric framework. All endeavours to measure fatigue represent an attempt to superimpose a continuum upon a psycho-physiological entity and as such, are capable of providing only a better or worse level of symbolic description (Hancock & Verwey, 1997). In the absence of an anchor scale (such as BAC in the case of alcohol or body temperature in the case of impairment due to thermal stressors), the multi-dimensional measurement of fatigue renders the concept susceptible to a certain amount of indeterminacy. Inevitably, ambiguity at the conceptual level creates related problems of data interpretation. This limitation is particularly striking during attempt s to measure the impact of multivariate fatigue on a complex, skilled behaviour such as driving.

The iterations within the trinity of stress postulated by Hancock & Warm (1989) render it difficult to make generic statements regarding how a defined task may impact on operator stress/fatigue or performance. The difficulty of this prediction is linked to the formulation of effort policy and its variable impact on performance effectiveness. It is suggested that this variability may account for the elusiveness of a consensual definition of operational fatigue and the development of any quantifiable and objective index.

The current study is concerned with the link between biological stressors and skilled performance as mitigated by mental effort policy. It is hypothesised that mental effort may be invested as means of protecting primary task performance against the influence of biological stressors (Hockey, 1993; Hockey, 1997; Mulder, 1986). Furthermore, the measurement of compensatory mental effort may constitute an appropriate basis on which to assess the influence of stressors on performance.

The effectiveness of compensatory mental effort may be assessed via the frequency and magnitude of those *latent decrements*, which may accompany degraded primary task performance (Hockey, 1993; Hockey, 1997). These decrements include: subsidiary task failure (i.e. a selective impairment of peripheral task components, both in the sense of physical localisation or low priority within the task hierarchy), strategic adjustment (i.e. a shift to routinised task strategies where mental effort is minimised), compensatory costs (i.e. increase of sympathetic nervous activity and related subjective appraisal), and fatigue after-effects (i.e. once mental effort has been invested for a sustained period, a subsequent increase of fatigue may reduce effort). Therefore, a consideration of latent decrements may reveal quantitative and qualitative patterns induced by stressors.

Mental effort was initially considered within the context of driving behaviour by Naatanen & Summala (1975). These authors suggested that effort investment was intimately associated with driver fatigue. Their principle claims are summarised as follows: (a) the necessary regulation of attention during driving is synonymous with effort investment, (b) a continuous process of effort investment leads to driver fatigue, and (c) that drivers may respond to rising fatigue by "trying harder" and/or adopting a cautious driving style, i.e. decreasing speed and increasing time headway to increase the safety margin. Evidence to justify the first hypothesis appeared in the years following the publication of this paper. De Waard (1991) found psychophysiological data (from analysis of 0.1Hz component of heart rate variability) indicating an increase of mental effort coinciding with a negotiation of highly demanding, roadway feature (i.e. an entrance to a high speed, highway). A similar investigation by Hancock, Wulf, Thom, & Fassnacht (1990) revealed a sporadic increase of workload and mental effort when drivers were involved in the negotiation of curves. A study of following behaviour by Fairclough, May, & Carter (1997) included a time of day manipulation, as a means of contrasting rush hour and off-peak levels of traffic density. Data from subjective workload and psychophysiology (the 0.1Hz component of heart rate variability) indicated a higher level of workload and mental effort investment during the rush-hour journey. These studies indicate that driving necessitates a dynamic investment of mental effort, in response to the ebb and flow of task demands in the roadway environment.

There is substantial evidence that a prolonged period of driving increases subjective estimates of fatigue and mental effort (Dureman & Boden, 1972; Fuller, 1984; Kecklund & Akerstedt, 1993; Riemersma & Biesta, 1978). In a study of professional drivers, Fairclough (1997) had subjects complete a 150 - 180 minute highway journey following a night shift. He found increased subjective effort over time, in line with parallel increases of inattention, boredom and sleepiness. In other words, subjects invested higher levels of mental effort to neutralise fatigue and discomfort. However, studies of prolonged driving using psychophysiological estimates of mental effort (i.e. mean Inter-Beat Interval (IBI) variability or 0.1Hz sinus arrhymia from the electrocardiogram) exhibited a different picture. For example, Brookhuis & De Waard (1993) found an increase of heart rate variability that was associated with increased time-on-task (indicative of decreasing levels of mental effort).

This adaptive mechanism of mental effort regulation places huge reliance on the accuracy of subjective assessment. The ability of the individual to monitor accurately external task feedback and the internal feedback from the self places an upper limit on the adaptability of mental effort policy. In addition, there is evidence that self-assessment may become increasingly inaccurate under conditions of high task demand and fatigue. For example, those operating conditions that produce dissociation between subjective estimates of workload and performance have been described by (Yeh & Wickens, 1988). This dissociation between appraisal and performance may be particularly marked for skilled performers (Baars, 1992; Langar & Imber, 1979). Brown (1994) and McDonald (1987) made a similar point concerning the influence of fatigue on self-monitoring of internal feedback. They claim that the detrimental effect of fatigue may extend beyond the impairment of performance, and degrade the quality of self-monitoring per se. Therefore, an impaired individual may be less able to assess accurately external and internal feedback, in order to make an appraisal of efficiency and mental effort may be misregulated as a consequence.

The appropriate regulation of mental effort hinges on the accuracy of feedback provided by self-monitoring and appraisal. The current study is intended to investigate how two different biological stressors (sleep deprivation and alcohol) influenced driving behaviour, latent decrements and self-appraisal.

5.3 Method

5.3.1 Design.

The study was designed to compare the effects of a between-subjects factor (treatment group) across two within-subjects factors (time-on-task and driving scenario).

The four treatment groups are described as follows: a *Control* group (subjects received a full night of sleep before the trial and did not receive alcohol), a *Partial Sleep Deprivation (PartSD)* group (subjects instructed to sleep for four hours between midnight and 0400 hours on the night before the trial and did not receive alcohol), a *Full Sleep Deprivation (FullSD)* group (subjects were instructed to remain awake throughout the night before the trial and did not receive alcohol), and the *Alcohol-Impaired (Alcohol)* group (subjects had a full night of sleep on the night before the trial and received an alcoholic drink (a mixture of vodka and lemonade) prior to the experimental session. The amount of alcohol administered was calculated based on the subject's body weight, according to the Widmark Equation (as described by Walls & Brownlie (1985)). Subjects received sufficient alcohol to approximate a peak Blood Alcohol Content (BAC) level in the range of 0.08 - 0.1 per cent (mg/g)).
The time-on-task associated with the journey was measured at three levels, each corresponding to a cumulative forty minute period within the two hour journey.

The driving scenarios encountered during the journey formed a second within-subjects factor. The scenarios were designed to induce differing degrees of task demand. Each forty-minute block was divided into six different scenarios and each scenario lasted for five minutes. The scenarios are described in the experimental task section.

5.3.2 Subjects

Sixty-four subjects participated in the experiment. All subjects were male and had normal or corrected 20/20 vision. Each subject was designated into one of four experimental groups, each containing sixteen subjects. The allocation of subjects into group was performed to balance the demographic variables of age, driving experience, annual mileage, driving frequency, average alcohol intake and average hours sleep per night across the four groups (Table 6). None of these demographic variables was statistically different between the four subject groups. On average, the subjects as a whole drove everyday, reported an annual mileage of greater than 10000 miles and less than 15000 miles, and slept for an average of 8 hours per night. All were paid for their participation.

SUBJECT GROUP	Age (yrs)	Driving Experience[1] (yrs)	Average Alcohol Intake[2]
CONTROL	30.63 (8.9)	12.8 (8.5)	16 (11.0)
PART SD	30.63 (9.0)	9.94 (5.60)	17 (9.1)
FULL SD	30.63 (9.8)	11.3 (10.1)	14 (12.0)
ALCOHOL	30.68 (9.8)	12.5 (9.0)	16 (10.5)

Table 6. Demographic characteristics of the four groups expressed as means and standard deviations (N=64). [1] number of years subjects held a full driving license, [2] self assessment of units of alcohol consumed per week.

5.3.3 Apparatus

The experiment was performed using a fixed-base driving simulator. Subjects were seated inside a Ford Scorpio and viewed a large projector screen (approximately 3 x 4m). A Pentium PC was used to simulate the vehicle model and to generate the driving scene. The computer-generated scene was projected onto the screen via a Sony Multiscan projector. The vehicle interior and the experimenter's location were linked via an intercom system. Electrocardiograph (ECG) data was collected via an analogue to digital converter (MacLab™) connected to a Macintosh Powerbook™ computer.

5.3.4 Experimental Task

The simulated driving scene consisted of a straight, two lane, left-side drive road flanked by road signs (which advised a 70mph speed limit), vegetation and marker posts. Off-road areas were coloured a uniform green on either side of the road boundary. The upper portion of the scene was coloured blue to emulate daytime driving. Vehicle dynamics included a small "wind gust" factor that was activated randomly throughout the session. Every alternate five second period, one of three outcomes occurred randomly: (a) no wind gust, (b) a wind gust approximating two steering degrees to the left sustained for five seconds and (c) a wind gust approximating two degrees to the right sustained for five seconds.

Unfortunately, the simulator software was not capable of providing interactive sound, but subjects were exposed to a recording of in-vehicle sound collected from a real vehicle travelling at an approximate speed range of 60-70mph on a motorway.

The subjects completed three forty-minute journeys in the course of an experimental session. Each block contained six driving scenarios as follows. The presentation order of the scenarios was fixed within each forty minute block, but different between the three blocks. The scenarios were as follows: *Open road* (the subject travelled in the left lane of an empty two-lane road that contained no other traffic), *Following* (the subject followed a lead vehicle in the left lane travelling at a steady speed of 60 mph), *Passing* (the subject travelled in the left lane whilst vehicles overtook in the right lane, the rate of vehicles passing the subject was approximately one every fifty seconds, *Following/Passing* (this scenario was a combination of both scenarios 2 and 3, i.e. the subject followed a vehicle in the left lane whilst vehicles overtook in the right lane), *Low sinusoidal following* (both sinusoidal following scenarios were based on an experimental manipulation reported by Brookhuis, De Waard, & Mulder (1994). In this scenario, the subject followed a lead vehicle in the left lane that systematically varied its speed sinusoidally between 55mph and 65 mph with a cycle time of 30 seconds), and *High sinusoidal following* (the lead vehicle again varied its speed sinusoidally. However, the amplitude of the speed change was twice that encountered in scenario 5, varying between 50mph to 70mph with the same cycle time.)

Each forty minute block began with a ten minute Open scenario in which no data was collected.

5.3.5 Experimental measures.

Three categories of experimental measures were used during the study: (a) primary vehicle control, (b) psychophysiology and (c) subjective measures.

<u>Primary Vehicle Control.</u>

Measurement of the driving task was divided into three categories: (a) lateral control (i.e. lane tracking), (b) longitudinal control (i.e. collision avoidance), and (c) speed control (i.e. a sub-task having substantial influence over lane tracking and collision avoidance).

Lateral control was measured at several levels of criticality: (1) frequency of "accidents" defined as those occasions when all four "wheels" left the left side lane edge, or when the subject either travelled across the right lane boundary and collided with a passing vehicle, or when the subject left the road on the right side lane edge, (2) frequency and mean duration of lane crossings when two vehicle wheels made contact with the left lane edge or the right lane boundary, and (3) frequency of incidents when the minimum Time-To-Line Crossing (TLC) (Godthelp, Milgram, & Blaauw, 1984) was 1 or 2 or 3 seconds. The calculation of minimum TLC (minTLC) excluded lane crossings. Thus, this classification scheme sub-divided lane tracking into a number of variables which varied along a continuum of severity. In order to index subjects' steering input to lateral control (the number of zero-crossings of the steering wheel as defined by McLean & Hoffman (1975) was measured.

Measures of longitudinal control or collision avoidance were only relevant to the sub-set of driving scenarios where a lead vehicle was present. The most critical (and obvious) indicator of impairment was the frequency of rear-end collisions. In addition, mean time headway (i.e. inter-vehicle distance divided by speed) was calculated.

All measures of primary driving performance were originally sampled at 10Hz and reduced to 2Hz for the purpose of analysis.

Psychophysiology.

ECG data was collected via three disposable electrodes attached to the subject's chest. R-peaks of the ECG trace were detected and corrected for artefacts. The resulting data was analysed with respect to mean Inter-Beat Interval (IBI) and IBI variability (heart rate variability, HRV). The HRV variable was subjected to spectral analysis, to decompose the HRV signal into three bandwidths (Mulder, 1979a; Mulder, 1979b; Mulder & Van Der Meulen, 1973): a low frequency band (0.02-0.06Hz) identified with the regulation of body temperature, a mid frequency band (0.07-0.14Hz) related to short-term blood pressure regulation and a high frequency band (0.15 - 0.50Hz) influenced by respiratory regulation. The mid-frequency band has been successfully employed to index mental workload in laboratory experiments (Aasman, Mulder, & Mulder, 1987) and during studies of driver mental workload (De Waard, 1996). In addition, analyses of the mid frequency band have yielded effects due to the influence of time-on-task (De Waard & Brookhuis, 1991; Mascord & Heath, 1992).

Subjective Measures.

The study employed a battery of subjective measurement questionnaires. Subjects completed multiple administrations of: (a) the UWIST Mood Adjective Scale (Matthews, Jones, & Chamberlain, 1990), (b) the NASA-Task Load Index calculated on the basis of raw ratings (RTLX) (Biers & McInerney, 1988; Byers, Bittner, & Hill, 1989; Hart & Staveland, 1988), (c) the Karolinska Sleepiness scale (Kecklund & Akerstedt, 1993), (d) a Cognitive Interference scale (Sarason, Sarason, Keefe, Hayes, & Shearin, 1986) and (e) an eight-point sobriety scale devised for the study (1= no noticeable effect, 2=slight increase of well-being, more relaxed, 3=definite increase of well-being and relaxation, 4=some minor impairment of judgement and reactions, 5=definite impairment of judgement and reactions, 6=significant impairment, 7=gross impairment, difficulty responding and 8=no longer able to drive).

5.3.6 Experimental procedure.

The subjects came to the Institute on two occasions, the first being a Practice session followed by a Test session on a different day. On most occasions, the Practice session took place in the morning prior to the day of the Test Session.

During the Practice session, subjects were weighed, completed a test of visual acuity and were fitted with ECG electrodes. Before the subjects entered the simulator, they received a set of standard instructions which emphasised the following points: (a) to stay in the left lane unless overtaking, and (b) to maintain a following distance of a maximum of three vehicles length to the lead vehicle and to keep the lead vehicle in sight at all times. The subjects performed a twenty-two minute "training" journey, which included short versions of the six driving scenarios.

Following the training journey, subjects completed the index of subjective questionnaires. The final phase of the Practice session involved a twenty-minute baseline journey to index subjects' "normal" behaviour in the simulator. The baseline condition contained only the Open Road and Steady Following scenarios (each scenario lasted for ten minutes), and primary vehicle performance and ECG data were collected. On completion of the baseline journey, subjects completed a second set of subjective questionnaires.

The two sleep deprivation subject groups were instructed to stay awake on the night before the Test Session. PartSD subjects were instructed to sleep for four hours between midnight and 0400, whereas FullSD subjects were instructed to remain awake during the whole of the night. To ensure subjects remained awake, they were required to leave a time-stamped message on a telephone answering machine once every sixty minutes, up until an hour before the Test Session.

All subjects were provided with transport to and from the Institute on the day of their Test Session. This arrangement was made to ensure the integrity of the alcohol placebo. The Test Sessions took place twice daily during the early afternoon (beginning after noon) and in the mid-afternoon from 15:00. The number of early afternoon and late afternoon sessions were equivalent among the four subject groups.

On arrival for the Test Session, all subjects were provided with a drink. The Control group and the two sleep-deprived groups received a placebo containing lemonade with a tablespoon of vodka floated on the surface to provide an alcoholic smell. In addition, the Alcohol group received specific instructions to eat a light meal at least an hour before the Test Session. Once the drink had been consumed, subjects were taken to the simulator where they completed a ten-minute "familiarisation" journey. Following familiarisation, all subjects completed a pre-test set of subjective questionnaires and were breathlysed using an alcometer (Lion Laboratories Alcometer S-D2). During the Test Session, the subjects were told that they would receive a modest financial punishment for every accident which occurred during the session (i.e. either if the car left the road or collided with another vehicle). This instruction was a deception intended to motivate the subjects to maintain performance (no subject actually suffered any financial penalties regardless of performance). Then the subjects completed the first forty-minute block of the simulated journey. On completion of the journey, subjects completed a second set of subjective questionnaires and were breathlysed once more. Both the subjective questionnaires and breathalyser tests were administered following completion of the second and third forty minute blocks. The recess between each journey block was restricted to five minutes duration. On completion of the Test Session, subjects were debriefed, paid for participation and transported to their homes.

5.4 Results

5.4.1 Alcohol Manipulation

Administration of alcohol using the Widmark equation is an inexact procedure and the means, standard errors and standard deviations of actual %BAC figures obtained during the study are shown in Figure

1. These data illustrate that the Alcohol subjects were on a descending BAC curve through the experimental session, dropping from a mean of 0.08% to 0.05% across the two hour journey.

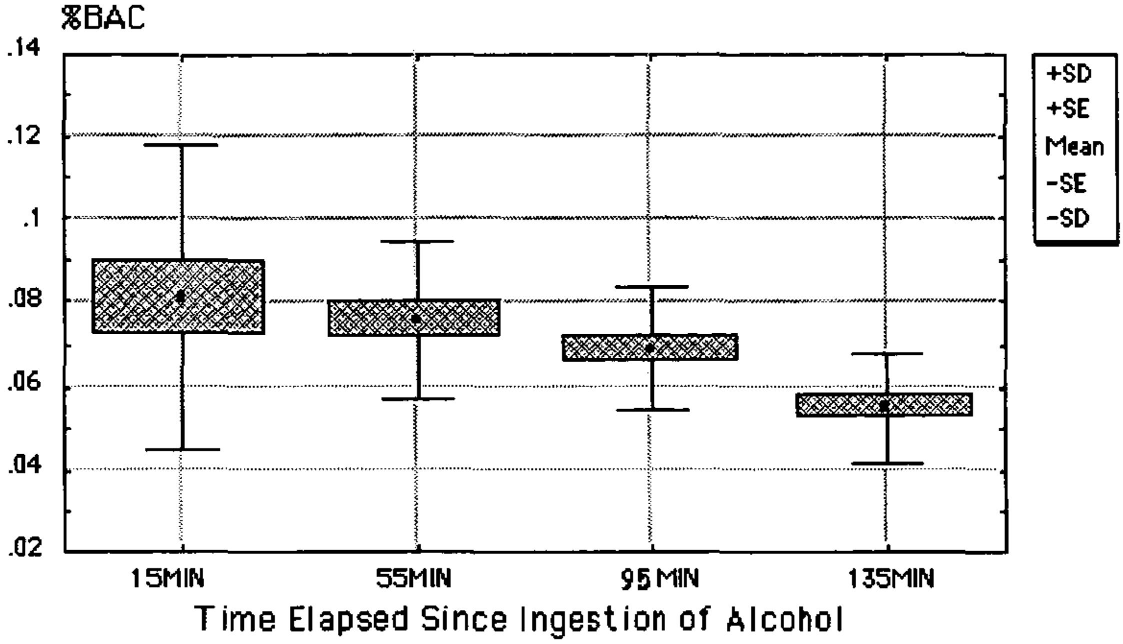


Figure 25. Box and whiskers chart illustrating the percentage Blood Alcohol Content levels recorded for the Alcohol group across the Test Session (N=16).

5.4.2 Primary task performance.

The raw vehicle data were averaged across a five-minute time window for each driving scenario and data were analysed using MANOVA techniques. When appropriate data were available from the baseline session and providing the test contained a maximum of three main effects, MANCOVAs were employed, i.e. using baseline data as a covariate. If the test contained more than 3 factors, this caused a substantial problem of interpretation, therefore the covariate was not included. Within-subject significance was reported using the Wilks Lambda [λ_w] statistic to overcome problems of sphericity (Vasey & Thayer, 1987). All data were tested for the existence of outliers (defined as raw values further than 2.5 standard deviations away from the mean). All post-hoc testing was performed via the Tukey HSD test.

Fifty-one "accidents" occurred in the course of the study. However, as only 10 subjects accounted for the total number of accidents, these data were not subjected to statistical testing.

The analysis of lateral control is presented with respect to two categories of data, those concerned with vehicle position and those capturing steering control. Descriptive statistics revealed the presence of an outlier in the lateral control data, therefore one subject from the Control group was not included in these analyses. All variables were subjected to a 4 (CONDITION) x 3 (TIME-ON-TASK) x 6 (SCENARIO) MANOVA. A significant main effect for CONDITION [$\underline{F}(3,59)=4.58$, $\underline{p}<0.01$] revealed that the frequency of lane crossings was higher for the FullSD group ($\underline{M}$ = 5.8 per 5 min) and the Alcohol group ($\underline{M}$ = 4.8), compared to the PartSD group ($\underline{M}$ = 3.4) or the Control group ($\underline{M}$ = 3.6). There was a significant effect of TIME-ON-TASK [$\underline{\lambda}_w(2,118)=0.76$, $\underline{p}<0.01$], i.e. the mean frequency of lane crossings rose by one per five minutes from the beginning to the end of the journey. The main effect of SCENARIO revealed that the highest frequency of lane crossings occurred during the Open and Passing scenarios [$\underline{\lambda}_w(5,295)=0.44$, $\underline{p}<0.01$]. A significant interaction between CONDITION x TIME-ON-TASK showed that the frequency of lane crossings was significantly lower in the Alcohol group compared to the FullSD group during the final 40 minutes of the journey [$\underline{\lambda}_w(6,118)=0.79$, $\underline{p}<0.01$].

The frequency of near-lane-crossings was indexed by those occasions when TLC fell below 2 seconds, i.e. the lateral velocity of the vehicle meant that it was 2 seconds away from a lane crossing. These data revealed a marginal effect due to CONDITION [$\underline{F}(3,59)=2.60$, $\underline{p}<0.06$] which indicated a higher frequency of near-lane-crossings for the PartSD group ($\underline{M}$ = 4.6 per 5 min) compared to the other three groups.

Steering control was analysed as the steering wheel reversal rate per minute. Multivariate analysis revealed a significant main effect for CONDITION [$\underline{F}(3,59)=26.6$, $\underline{p}<0.05$], indicating that mean steering wheel reversal rate was higher for the Control and Alcohol groups ($\underline{M}$ = 15.6 and 14.3 respectively) compared to the PartSD and FullSD groups ($\underline{M}$ = 11.2 and 10.9 respectively). A significant main effect for SCENARIO [$\underline{\lambda}_w(5,295)=0.56$, $\underline{p}<0.01$] revealed that levels of steering activity were higher during the High Sinusoidal and Following scenarios ($\underline{M}$ = 15.0 in both cases) than the Open scenario ($\underline{M}$ = 12.6).

The main effects for all three measures of lateral control across the four experimental groups are represented in Figure 26.

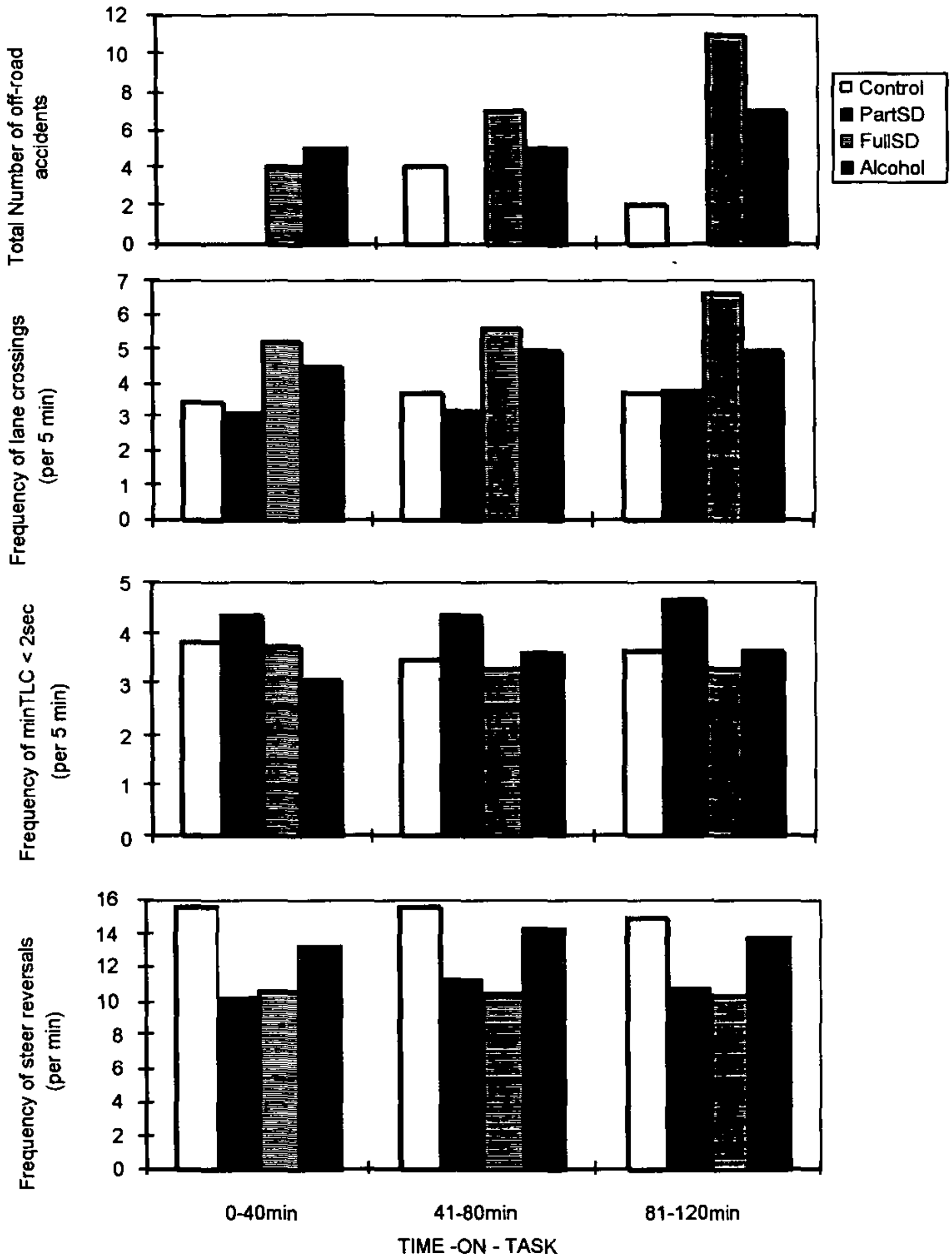


Figure 26. Lateral vehicle control as represented by: mean frequency of lane crossings, mean frequency of near lane crossings and steering wheel reversal rate across the four treatment groups (N=63).

Analysis of mean time headway was confined to those four scenarios in which a lead vehicle was present. Six subjects were omitted from this analysis due to the occurrence of accidents or merely losing sight of the lead vehicle. This analysis showed a marginal effect for CONDITION [$\underline{F}$ (3,54) = 2.22, $\underline{p}$ =0.08]. Post-hoc tests revealed a significant difference between the Alcohol ($\underline{M}$ = 3.10s) and FullSD groups ($\underline{M}$ = 3.98s). There was also a significant effect for TIME-ON-TASK [$\underline{\lambda}_w$(2,108) = 0.88, $\underline{p}$ < 0.05]; mean headway decreased by 0.5s between the initial 40 minutes and the final 40 minutes of the journey.

Speed variability was measured as standard deviation of speed (across a 5 minute time window). The analysis showed a significant main effect for CONDITION [$\underline{F}$ (3,59) = 4.56, $\underline{p}$ < 0.01]. Post-hoc testing revealed that speed variability was lower for the Alcohol group ($\underline{M}$ = 4.2 mph) than either the PartSD or FullSD groups ($\underline{M}$ = 5.7 mph for both). In addition, there was a significant main effect for TIME-ON-TASK [$\underline{\lambda}_w$ (2,120) = 0.89, $\underline{p}$ < 0.05] providing evidence of an increase in speed variability over time.

5.4.3 Psychophysiology

It is known that alcohol has a confounding effect on mean IBI and the 0.1Hz component of HRV (Gonzalez-Gonzalaz, Llorens, Novoa, & Valeriano, 1992). Therefore, ECG data from the subjects in the Alcohol group were not subjected to statistical testing. In addition, ECG data was lost from 7 subjects across the three remaining subject groups due to measurement artefacts.

The raw IBI data was subjected to analysis using CARSPAN software (Mulder & Schweizer, 1993) to isolate the mid-frequency component of heart rate variability. This data was subjected to a natural log transform and a baseline conversion prior to parametric MANCOVA analysis as described by (Meijman, 1995).

There were significant main effects for CONDITION [$\underline{F}$ (2,32)=2.52, $\underline{p}$ =0.05] and TIME-ON-TASK [$\underline{\lambda}_w$(2,64) = 0.64, $\underline{p}$ < 0.01]. It was apparent that the mid-frequency component was significantly suppressed for the PartSD group compared to the Control group (p<0.05). In addition, mean power in the mid-range frequency increased between the first and third period of the journey (p<0.05). These findings are indicative of a higher level of mental effort investment for the PartSD group compared to Control subjects. This main effect is illustrated in Figure 27.

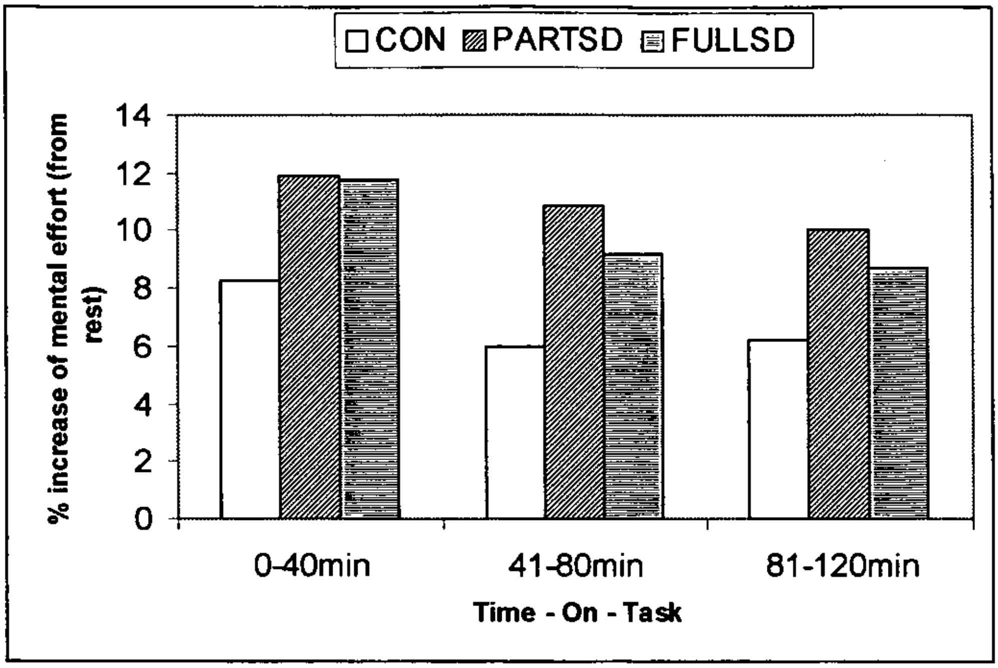


Figure 27. Mean power in the mid-frequency bandwidth of heart rate variability for three experimental groups across the test session (N=41).

5.4.4 Subjective Data.

Subjects completed the questionnaires on four occasions through the Test session, including a pre-test administration. The subjective data were analysed in a series of 4 (CONDITION) x 4 (TIME-ON-TASK) MANCOVAs with data collected from the end of the practice session used as a covariate. The results of these analyses are shown in Table 2 (it should be noted that there were no significant interaction effects).

		CONDITION		TIME ON TASK	
		F(3,60)	P	λ_w (2,120)	P
MOOD	Energetic	5.88	<0.01	0.54	<0.01
	arousal	1.43	NS	0.94	NS
	Tense arousal	1.34	NS	0.65	<0.01
	Hedonic Tone				
WORKLOAD	Mental demand	1.09	NS	0.92	NS
	Physical demand	4.41	<0.01	0.5	<0.01
	Time demand	4.06	<0.05	0.95	NS
	Performance	2.73	0.05	0.80	<0.01
	Effort	3.16	<0.05	0.87	<0.05
	Frustration	4.51	<0.01	0.61	<0.01
	Mean workload	2.63	0.05	0.68	<0.01
COGNITIVE	Task Relevant	4.23	<0.01	0.60	<0.01
INTERFERENCE	Task Irrelevant	1.29	NS	0.45	<0.01
SLEEPINESS	Karolinska scale	11.37	<0.01	0.39	<0.01
SOBRIETY	Self-Rating	12.11	<0.01	0.81	<0.05

Table 7. Results of the MANCOVA analyses of subjective scales (N=64).

Measurement of subjective mood revealed a significant decline of energetic arousal for the two sleep-deprived groups compared to Control subjects. In addition, sleep deprivation raised subjective workload via increased effort and reduced estimates of performance efficacy. The FullSD group only experienced increased temporal demand and physical demand. The effect of Alcohol was to reduce the level of frustration experienced by subjects, as compared to the other three groups. It was found that the frequency of task relevant thoughts was higher for the FullSD group than either the Control or Alcohol groups. The PartSD group exhibited a similar pattern, but frequency was only significantly higher than the Control group. Both subjective ratings of sleepiness and sobriety were included to reference the experimental manipulations. The mean values for these scales are shown in Figure 28. Subjective sleepiness was significantly higher for both sleep-deprived groups compared to the Control group (but did not differentiate between PartSD and FullSD). Self-rated drunkenness was highest for the Alcohol group.

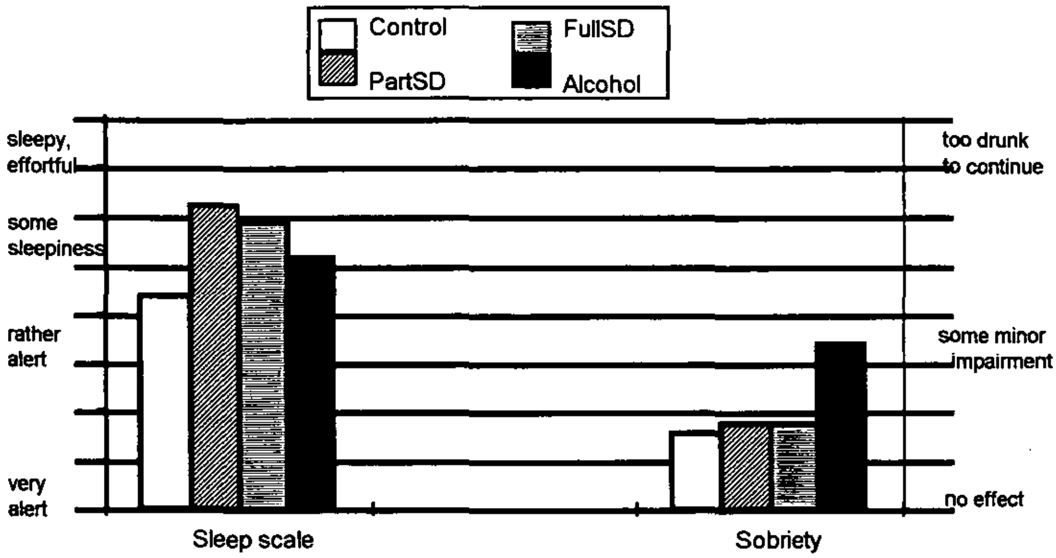


Figure 28. Mean values for subjective ratings of sleepiness and sobriety for all four subject groups (N=64).

The TIME-ON-TASK effects for subjective measures were all in the expected direction, i.e. workload factors, task-irrelevant thoughts and subjective sleepiness all increased across the trial duration; whereas energetic arousal, hedonic tone and task-relevant thoughts all decreased.

5.4.5 Performance Efficiency Index

The analysis of performance efficiency involves the interaction between performance effectiveness and psychophysiological mental effort (Meijman, 1995). In this case, it was not possible to obtain appropriate psychophysiological data from those subjects who had received alcohol. Performance efficiency is represented in Figure 29 by indexing psychophysiological mental effort (subjected to the % baselining procedure described by Meijman) against the frequency of right-side lane crossings per 5 minutes.

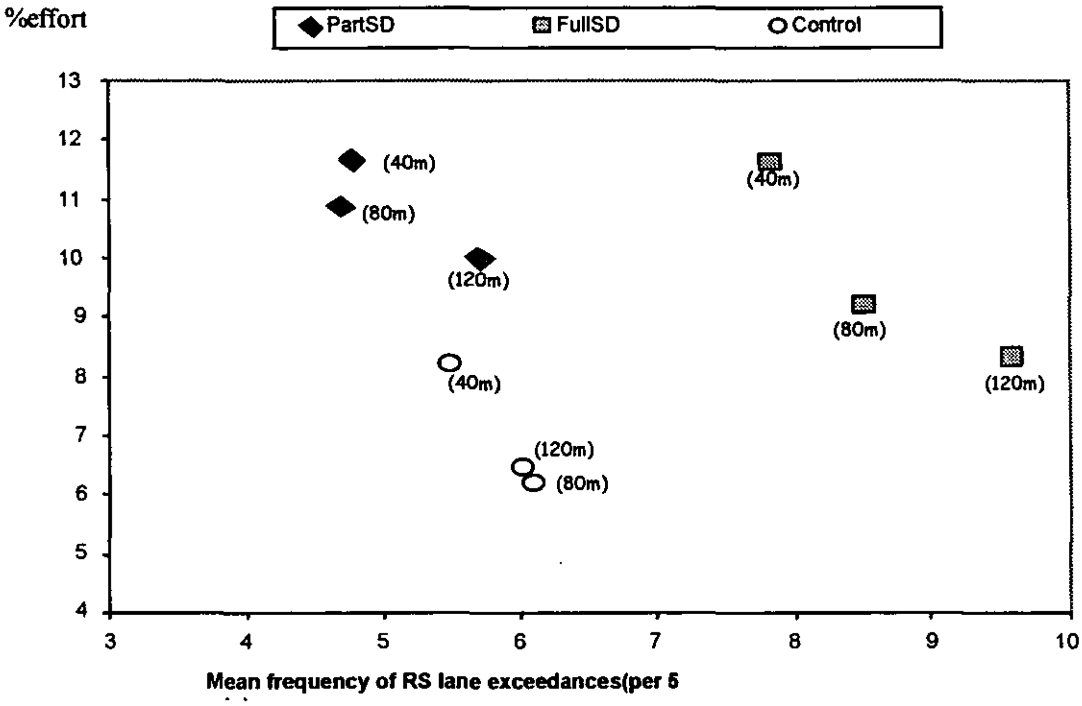


Figure 29. Performance efficiency index for three experimental groups across time-on-task (N=48). NB: increase of x-values = poor lateral control of vehicle, increase of y-values = increased mental effort.

The analysis of performance efficiency revealed that the accuracy of lateral control exhibited by the PartSD subjects was equal to, and even slightly superior to the Control subjects. However, it was significant that the PartSD group had to exert a substantially higher level of mental effort in order to achieve equivalent performance to the Control subjects. The FullSD subjects also invest a high level of mental effort during the initial period of the simulated journey. Despite this trend, the FullSD subjects were unable to improve the accuracy of lateral control. Once the FullSD group passed the 80min mark, they were unable to sustain mental effort and performance effectiveness was observed to decline accordingly.

5.5 Discussion

The effects of sleep deprivation and alcohol on driving performance may be categorised in terms of safety critical changes (i.e. driver errors likely to increase the probability of accident) and non-safety critical changes (i.e. alterations in vehicular control which do not increase the probability of accident).

It was apparent that a full night without sleep was sufficient to impair those subjects with respect to both categories. The FullSD group exhibited the highest frequency of lane crossings (Figure 26) and the highest mean time headway separation to a lead vehicle. The former effect may have been the direct result of a reduced level of steering input, whereas the latter may have been an adaptive strategy to counteract an increased risk of rear-end collision.

The effects of a reduction of sleep were more paradoxical in the case of the PartSD group. These subjects exhibited normal lateral control (relative to the Control group), whilst functioning on the reduced level of steering input which characterised the FullSD group (Figure 26). The only evidence of impairment for the PartSD group was an increased frequency of near-lane-crossings (Figure 26). These data suggest that PartSD subjects exhibited a proactive strategy of lateral control. This strategy was described with reference to the categorisation of latent decrements described by (Hockey, 1993; Hockey, 1997).

The types of driver errors induced by functional impairment were designated with respect to criticality. This analysis of lane-keeping behaviour is illustrated in Figure 26. Alcohol and full sleep deprivation manipulation produced the highest number of accidents and lane crossings. The PartSD group committed no accidents and was indistinguishable from the Control group in terms of the number of lane crossings observed. However, the analysis of minTLC revealed a higher number of near-crossings for the PartSD group. This pattern revealed a subsidiary degradation of lane-keeping with respect to near-crossings for the PartSD subjects, which did not produce either a higher number of lane crossings or accidents. This finding is presented as an example of subsidiary task failure (Hockey, 1993; Hockey, 1997). In this case, it is hypothesised that lane-keeping performance was allowed to degrade to a pre-critical point prior to error-correction. This pattern of performance is similar to the widening of the indifference range due to fatigue reported by Bartlett, 1953; 1943).

A similar effect was observed with respect to lane crossing errors. Given that the right lane was devoid of vehicles for the majority of the journey, the risk of an off-road accident was greater for crossings the left-side road boundary (as opposed to a right-side lane boundary). An analysis of lane crossing frequency revealed that crossing at the right-side boundary occurred more frequently than the left-side road boundary (respective means = 6.6 and 2.2 per 5 minutes). In addition, crossings on the right-side boundary were, on average, over twice as long as those on left-side road edge (respective means = 11.5

and 5.4 seconds). Therefore, the degradation of simulated driving occurred in a manner that minimised the risk of accident. This finding illustrated how risk perception plays a role in the utility assessment underlying subsidiary task failure. The driver prioritises sub-goals and performance standards in a rational manner. This process determines which error forms are less-critical and therefore constitute "acceptable" categories of performance degradation.

In contrast to their sleep-deprived counterparts, the impairment of performance exhibited by the Alcohol subjects was limited to safety-critical changes such as more frequent lane crossings (relative to Control and PartSD subjects) and a reduction of mean time headway (relative to the FullSD group). A second feature which distinguished the FullSD and Alcohol subjects was an interaction with time-on-task. The results revealed that impairment due to sleep deprivation tended to increase with time-on-task, i.e. lane crossing frequency peaked for the FullSD group during the final forty minutes. By contrast, the impaired lateral control due to alcohol was stable across the journey despite a descending alcohol absorption curve (Figure 26).

The latent decrement described by Hockey (1993;1997) as strategic adjustment represents a conservation strategy to minimise mental effort investment. However, this does not necessarily imply a primary performance decrement or even a subsidiary task failure. The assessment of strategic adjustment focuses on the efficiency of task performance (Eysenck & Calvo, 1992). For example, consider the data in Figure 26 as a bottom-up chain of control input, i.e. from input as steering wheel activity to various errors of lateral control. The patterning of data (considered as a whole) as relative to the Control group is quite striking. The FullSD subjects have the highest number of accidents and lane crossings, coupled with a low level of steering activity (indexed by reversal rate). Therefore, a low rate of steering input accounted for the high number of errors committed by the FullSD group. The Alcohol group has a higher number of accidents and lane crossings comparative to both the PartSD group and the Control group. However, the Alcohol group has sustained a level of steering wheel input that is equivalent to the Control group. The implication here is that Alcohol subjects performed highly inefficiently with respect to lateral control , i.e. high amount of input coupled with a low quality of lateral control performance. By contrast, the PartSD group revealed a very efficient level of lateral control. These subjects provided a low level of steering input equivalent to the FullSD group, but were capable of lateral control as accurate as the Control group. A pattern of reduced task activity characterised mental effort conservation and was apparent in both sleep-deprived groups. The difference between the groups was that the PartSD group was capable of an effective and efficient performance strategy.

An analysis of time headway indicated that FullSD subjects increased the following distance to a lead vehicle, whereas the Alcohol group exhibited a slight decrease of time headway. These changes are characteristic of an adaptive response at the tactical level in response to an awareness of falling standards of operational performance for the FullSD group (Van Winsum, 1996). The implication was that awareness of poor performance caused a strategic increase of the safety margin for the FullSD group. The response of the Alcohol group was to reduce the safety margin, thereby illustrating a characteristic lack of sensitivity at the operational level (Van Winsum, 1996). These findings illustrated the importance of external feedback for adaptive driver behaviour.

The six driving scenarios were included to induce a variable level of mental workload throughout the simulated journey. The presence of a lead vehicle was the crucial feature, which differentiated the level of lateral control across all groups. The absence of a lead vehicle during Open and Passing scenarios caused an increased frequency of lane crossings. The improvement of lateral control associated with a lead vehicle may have been due to either (a) a perceptual effect (i.e. the presence of vehicle on the simulated view functioning as a perceptual cue), or (b) a workload effect (i.e. the lead vehicle functioning as a potential hazard, and demanding a higher level of attention and control), or a combination of both effects. The analysis of psychophysiology provided no evidence of increased mental effort when a lead vehicle was present. However, steering wheel reversal rate was elevated for scenarios which included a lead vehicle (e.g. Following, High Sinusoidal) compared to the Open scenario. Therefore, the presence of the lead vehicle appeared sufficient to stimulate steering activity, if not to raise mental effort.

It is postulated that any compensatory response to sleep deprivation was triggered by increased subjective discomfort and an awareness of reduced performance efficacy, which accompanied operational fatigue (Table 2 and Figure 28). These subjects responded by mobilising subjective effort to counteract the influence of sleep deprivation. This response was demonstrated by the analyses of subjective workload ratings (i.e. increased effort, increased frequency of task-relevant thoughts - Table 7).

A strategy of mental effort compensation was apparent from the psychophysiological data, where the PartSD group exhibited higher mental effort relative to the Control subjects (Figure 27). The psychophysiological data analysis suggested that PartSD subjects were capable of sustaining the higher level of mental effort necessary to counteract sleepiness throughout the simulated journey, whereas the FullSD group could not. This interpretation was supported by the analysis of performance efficiency (Figure 29). This index clearly indicates that PartSD subjects exerted high levels of mental effort in order to achieve equivalent performance to the Control subjects. This pattern represents an effective but inefficient strategy of mental effort investment, i.e. adequate performance effectiveness obtained but only via increased effort investment. On the other hand, the FullSD group initiates a high level of mental effort investment but are unable to improve performance. The data from the FullSD group represents a strategy of mental effort investment that is both ineffective and inefficient.

It was apparent from the analysis of subjective data that a differential awareness of subjective discomfort was one feature that distinguished the sleep-deprived subjects from the Alcohol group. Aside from decreased sobriety, the Alcohol subjects exhibited reduced frustration comparative to the other treatment groups (Table 7). It is hypothesised that this effect originated from the increase of well-being induced by alcohol. In addition, the Alcohol subjects rated subjective estimates of performance to be superior to the two sleep-deprived groups across the simulated journey as a whole (Table 7), when performance measures indicated that this group performed as poorly as the FullSD subjects (Figure 26). It is known that alcohol tends to promote an inflated sense of confidence (Walls & Brownlie, 1985) and this factor may have accounted for the distorted perception of performance that typified the Alcohol subjects.

It was unfortunate that it was not possible to obtain psychophysiological measures of mental effort or to include the Alcohol subjects within the performance efficiency index. Based on the model of mental effort regulation (Chapter 4), it would be anticipated that any stressor such as alcohol, which is capable of distorting both internal and external sources of feedback, would completely disable the process of effort regulation. The current analyses illustrated that the Alcohol subjects perceived lower levels of discomfort (relative to all groups) and higher performance effectiveness (relative to the sleep-deprived subjects). Therefore, there was no increase of subjective effort because these subjects did not perceive either an internal or an external spur for effort investment. However, this hypothesis was only supported by subjective measures within the current study.

The analysis of subjective data shown in Table 7 illustrates a level of equivalence between the two groups of sleep-deprived subjects. Both groups perceived increased sleepiness, reduced energy, declining performance and increasing levels of frustration. In addition, both groups attempted to counteract these detrimental influences by increasing subjective effort and the frequency of task-relevant thoughts. Despite these similarities, the PartSD group was able to sustain mental effort and primary task performance whereas the FullSD subjects could not.

There were several points of dissociation between the two sleep-deprived groups and these deviations may indicate the differences between an effective and an ineffective strategy of effort investment. First, the FullSD group perceived both physical demand and temporal demand sub-scales of the TLX to increase whereas the PartSD group did not (Table 7). The increase of physical demand may have indicated accelerated symptoms of muscular fatigue induced by continuous steering input (Sheridan, et al., 1991). It was interesting that FullSD subjects perceived the temporal demands of the task to be relatively accelerated compared to other groups. This finding may be interpreted as evidence of increased cognitive inertia due to sleep deprivation, e.g. increased frequency of 'blocking' (Bills, 1931), slowing of overt responses (Dinges & Kribbs, 1991) – which may provide the illusionary impression that task pacing is accelerated. Both findings illustrate that the symptoms of sleepiness are more pervasive in the case of the FullSD subjects relative to the PartSD group. This distinction suggest that our perceptions of our capability to sustain performance or 'coping capacity' (Schonpflug, 1983) or Distress (Matthews, et al., 1997) may be at the root of an effective effort investment strategy.

5.6 Conclusions

The study contrasted the influence of different qualitative stressors (alcohol, sleep deprivation) as well as the quantitative influence of partial- and full-sleep deprivation on driving behaviour. It was apparent that a full night without sleep degraded subjects' lane-keeping performance to approximately the same extent as an illegal amount of alcohol. However, this quantitative similarity was superficial as degraded performance originated from different causal factors. In the case of the sleep-deprived subjects, these individuals were aware of degraded performance, increased sleepiness, and therefore, attempted to invest effort to protect performance. This strategy proved ineffective and consequently, these subjects incurred costs in form of heightened frustration. The Alcohol group did not perceive performance to degrade and experienced an increased sense of wellbeing as was evident from their low level of frustration. This comparison indicated the significance of internal and external feedback cues as inputs to effort regulation strategy. If these cues are not available, an appropriate strategy to counteract the influence of stressors on performance may not be possible.

The comparison between sleep-deprivation conditions illustrated the distinction between successful and unsuccessful effort strategies. The PartSD group was capable of sustaining equivalent lateral control performance to the Control subjects by increased psychophysiological effort. In addition, these subjects operated at a lower level of steering input and made more frequent near-errors, i.e. their level of vehicular control was the most efficient of all four groups. It is hypothesised that a combination of increased mental effort and efficient vehicular control allowed these subjects to sustain performance throughout the task.

6 THE INFLUENCE OF PERFORMANCE FEEDBACK ON DRIVER BEHAVIOUR[10].

6.1 Abstract

Eighteen male subjects were requested to perform a simulated driving journey until "they felt unable to continue at an acceptable level of performance." Financial incentives were offered for time spent "driving" and subjects were informed that they would be financially punished if they "crashed" the driving simulator. The subjects performed this task under three conditions: (a) in the presence of a Discrete Display where subjects were provided with discrete impairment feedback, (b) a Continuous Display condition where subjects were provided with continuous impairment feedback, and (c) a Control condition where no feedback was made available. Feedback was based on an on-line analysis of steering control and lateral control, and was updated every 30 seconds. Both displays provided three levels of feedback, i.e. normal (<=40% impairment), minor impairment (41-70% impairment) and major impairment (>71% impairment). It was found that the presence of feedback had no effect on average duration of the simulated journey, and exerted little influence over the decision whether or not to terminate the journey. However, subjects produced a higher level of lateral accuracy when provided with warning feedback. There was no evidence that feedback presence increased the level of mental effort, exerted any influence on mood or subjective fatigue, or increased subjective workload. It is concluded that Feedback affected the qualitative investment of mental effort without increasing the quantity of effort, i.e. performance effectiveness was increased without additional effort investment.

6.2 Introduction

A survey of drivers in the UK conducted by Maycock (1995) revealed that 29 per cent of his survey sample (N=4600) had felt close to falling asleep at the wheel during the previous twelve months. However, only 7 per cent of the same sample reported fatigue as a contributory factor in accident involvement (adjusted by Maycock to implicate fatigue in between 9 and 10 per cent of accidents for the sample). The general point, which may be gleaned from this discrepancy, is that an extreme experience of fatigue does not have a one-to-one correspondence with accident involvement. This is

[10] A version of this study has been submitted for review in Transportation Human Factors.

unsurprising in itself. Individual drivers may employ various strategies during naturalistic driving to counter the influence of fatigue which may be employed at more or less regular intervals, e.g. open car window for fresh air, stop the vehicle and take a walk.

In his writing on the topic of driver fatigue, Brown (1994) revised the perspective of earlier theorists (Bartley & Chute, 1947), which characterised operational fatigue as a subjective disinclination to continue performance. It is presumed that this 'subjective disinclination' increases as a function of time-on-task and serves to provide feedback to the individual regarding the magnitude of fatigue. Once a driver perceives their disinclination to rise above a critical level, they may decide to take a break from the driving task. However, there may also be a number of reasons for the driver to ignore the accumulation of driver fatigue and remain on the road, e.g. lack of suitable stopping place, tight journey schedule, desire to reach destination.

The decision whether or not to break from the driving task may be viewed within an adaptive framework originally postulated by Cameron (1973) and advanced by (Hancock & Warm, 1989). The latter authors claimed that the act of sustaining attention may be treated as a source of stress in its own right hinging on three iterative foci: the input stress (i.e. task demand, sleep deprivation, noise), the adaptive or compensatory response to input stress and the resultant influence on the goal-directed outputs of the individual. Within this conceptualisation, the input stress of a 2-4 hour car journey would be identical for all drivers. However, individual differences would be apparent when a range of more or less successful compensatory strategies were employed. A successful strategy will achieve two related goals related to performance efficiency (Eysenck & Calvo, 1992; Schonpflug, 1985), i.e. to sustain an adequate level of performance whilst minimising mental effort expenditure and associated costs such as stress or discomfort.

The previous chapter contrasted the influence of various stressors on driving performance. This study concluded that the accuracy of external feedback was crucial for mental effort regulation. Those subjects deprived of sleep exhibited a higher awareness of performance impairment than intoxicated subjects. This awareness had a knock-on effect for the appraisal of internal costs.

Little research has been performed to investigate the role of external feedback within the domain of driver fatigue. There is some evidence that drivers have a tendency to overestimate their own capabilities (Brown, 1990). This perspective is supported by research on skill development and awareness (Baars, 1992; Langar & Imber, 1979), i.e. a high level of skill is equated with a low level of awareness of one's own performance.

Both McDonald (1987) and Brown (1994) have advanced the hypothesis that rising fatigue may reduce the individual's capacity to self-monitor, with respect to both declining ability or the experience of fatigue. According to McDonald ".. anything that impairs self-monitoring, particularly in terms of over-estimating one's psychological resources or under-estimating the demands to be met, is likely to give rise to unpredictable or uncontrollable effects of fatigue" (p. 187). This hypothesis raises a number of interesting questions for fatigue and the regulation of mental effort. In both cases, fatigue may exert a bias on the regulatory mechanism, such that we tend to over-estimate our performance capacity or disregard internal feedback regarding symptoms of fatigue or stress. There are several natural outcome of degraded effort regulation: (a) effort investment is misdirected, i.e. non-critical aspects of performance are preserved at the expense of critical aspects, (b) effort policy is inefficient, i.e. effort invested at a much higher level than is necessary to sustain adequate performance effectiveness, and (c) finite effort reserves are over-extended, i.e. individual overestimates ability to sustain performance over a given period and is forced to adopt a strategy of effort conservation, regardless of its impact on performance.

One countermeasure to the problems of ineffective effort regulation due to distorted external feedback is the design and development of telematic systems to monitor driver behaviour. These systems encompass real-time monitoring, diagnosis and feedback of driving impairment in various forms (Brookhuis, De Waard, & Bekiaris, 1997; Fairclough, 1997; Fairclough & Hirst, 1993; Haworth & Vulcan, 1991; Mackie & Wylie, 1990; Wierwille, 1994). One goal of these systems is to provide a source of objective feedback to counteract the influence of fatigue on self-monitoring. However, like their human counterparts, these diagnostic systems are reliant on the sensitivity and validity of sensor apparatus.

The measurement of driving performance represents a valid strategy to index impairment (i.e. impairment is inferred directly based on primary task performance rather than via a proxy measure). However, there are a number of problems associated with the use of driving performance as a predictive and diagnostic source of impairment (Fairclough, 1997). Specifically, primary task measures are deemed to be insufficiently sensitive to the presence of impairment, i.e. task performance is "protected" from the influence of impairment by compensatory strategies (De Waard, 1996; Hockey, 1997). Despite these problems, primary task measures occupy a pivotal role in the development of driver impairment monitoring systems. These measures form the main source of data for those telematic systems that employ multidimensional assessment (Brookhuis, De Waard, & Bekiaris, 1997). In addition, the standard means of estimating the validity of indirect measures such as psychophysiology is to utilise changes in primary driving performance as a reference variable (Wierwille, Wreggit, & Knipling, 1994).

The aim of the current study was to investigate the influence of performance feedback on driver behaviour. A secondary goal of the study was to explore how alternative forms of interface designs may exert an influence driver behaviour, i.e. the provision of discrete vs. continuous performance feedback. In addition, it was important to consider any influence on fatigue accumulation and mental workload, which may be exerted by the alternative forms of warning display.

6.3 Method

6.3.1 Experimental Design

The experiment was designed as a repeated measures experiment. All subjects participated in three sessions: (a) a Control condition (no feedback present), (b) the Discrete warning system condition, and (c) the Continuous warning system condition. The order of presentation of each condition was counterbalanced across the subjects.

The experimental sessions were scheduled for three times during the day (10:00, 12:30, 15:00). Each subject was scheduled for the same time-period for each experimental session. There was at least a 48 hour period between each consecutive session for all subjects.

6.3.2 Simulated impairment warning system

The simulation of a monitoring system was based on two performance-based indicators of driver impairment, these were:

- The standard deviation of vehicle lateral position
- The proportion of high velocity steering corrections

The former measure has been validated via numerous studies of both on-road and simulator-based studies of driver impairment (Brookhuis & De Waard, 1993; Riemersma, Sanders, Wildervanck, & Gaillard, 1977). The latter measure corresponds to the upper 20th percentile of steering velocity as a proportion of a "baseline" period representative of "normative" driving, i.e. the number of fast steering corrections, which may be indicative of a sudden corrective movement to prevent the vehicle straying from the lane. Relative values of both criterion measures were calculated with reference to baseline values collected during a ten minute familiarisation session on the simulator. Criterion measures were calculated every 30 seconds according to the following formulae and used as input to the simulated warning interface. These criterion were enforced with reference to the highest value on an EITHER/OR basis, i.e. feedback was triggered with respect to the highest of the two criterion measures.

LPcrit = LPsd (*30s*) / baseline LPsd

STVEL = % of data (*30s*) < 80^{th} percentile value of baseline / 20

The simulated warning system appeared in two forms during the experiment: (1) as a discrete feedback system, and (2) as a continuous feedback system. In both cases, three levels of performance feedback were provided which were colour-coded into green, amber and red zones. The green zone was defined as 'normative' performance, the amber zone signified 'moderate impairment' and the red zone indicated 'severe impairment'. These zones were quantified by the percentage change of either Lpcrit or STVEL relative to baseline values as shown in Table 8.

Impairment Zone	Quantification by criterion values
GREEN	<=40 % above baseline values
AMBER	>41% and <=70% above baseline values
RED	>71% above baseline values

Table 8. Quantification of warning feedback

The discrete warning display (DD) corresponded to three coloured lights (FIG). Under normative circumstances, only the green light was activated. The Continuous warning display (CD) corresponded to a dial with three coloured sections (FIG). In this case, each of three impairment zones was divided into three sub-sections and the pointer moved within, as well as between each coloured section. In both cases, each display was updated every 30 seconds.

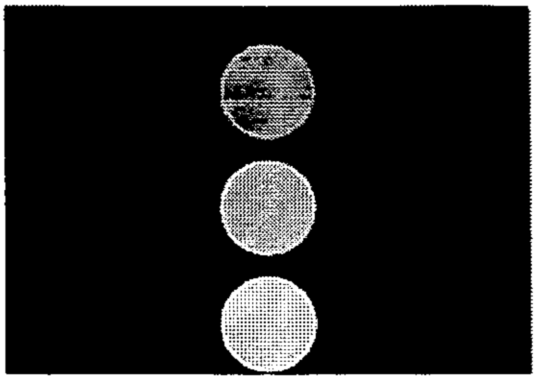
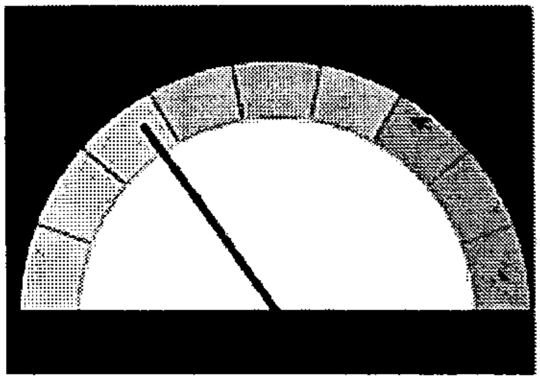



Figure 30. Schematic representations of the discrete and continuous warning interfaces.

Both versions of the warning interface were supplemented by a spoken warning, which sounded if subjects exhibited moderate or severe impairment. These warnings were as follows:
Amber warning: "Warning. You are showing symptoms of impairment."
Red warning: "You are highly impaired. Please take a break."
These voice warnings were activated during once during every 30 second period that the subject remained in either the amber or the red zone.

6.3.3 Driving simulator

The driving task was simulated via PC-based computer software devised at the TNO Institute for Human Factors. This software simulated a colour view of a single-vehicle road on a large 20" monitor. The roadway geometry in the simulated world represented a large circuit of straight road sections interspersed with curved sections to the left and to the right. The software also provided a speedometer and engine sound feedback. If the simulated vehicle drifted from the grey road area to off-road green area, subjects received noise feedback to aid error-correction. The simulated roadway circuit contained a "Services" sign at the mid-point. This sign functioned as a cue for subjects to elicit subjective data or to terminate the journey (see Procedure section below). Vehicular control input was achieved via adapted version of the ThrustMaster™ steering console with pedals for accelerator and brake control.

6.3.4 Apparatus

The driving task was simulated via Pentium computer software with the screen placed at eye-level. The simulator PC was linked to a second Pentium via an Ethernet connection. This PC contained software written specifically for the experiment, capable of converting real-time, raw data from the simulator into baselined criterion values. These values were transmitted to a third PC that contained the necessary software to receive data and to simulate warning feedback. The simulation of each warning interface was achieved via software created in VisualBasic5.0™ specifically for the experiment.

ECG data was collected via an analogue-to-digital converter (MacLab™) working in conjunction with specific software (Chart™) on a Macintosh Powerbook.

6.3.5 Subjects

18 male subjects participated in the study. The average age of the subjects was 34.2 years (s.d. = 9.9) and they had been fully qualified drivers for an average of 16.5 years (s.d. = 10.7). The majority of the drivers used their vehicle everyday and most approximated their annual mileage to fall between 9000 and 24,000 km.

6.3.6 Experimental Measures

<u>Primary task performance</u> was indexed by a number of variables related to the quality of driving performance. These measures were as follows:

- Average value of LPcrit
- Average value of STVEL
- Time spent in the Amber zone (as a fraction of total journey duration)
- Time spent in the Red zone (as a fraction of total journey duration)
- Mean frequency of lane crossings (per minute), i.e. when either right or left wheel of simulated vehicle made contact with lane boundary
- Mean frequency of near-crossings (per minute), i.e. when either right or left wheel was 20 cm or less from lane boundary
- Mean steering wheel standard deviation

• Mean speed (Km/h)

Psychophysiological measures constituted the electrocardiogram trace (ECG) recorded at 1000Hz. The subsequent inter-beat interval (IBI) was subjected to power spectrum analysis in order to distinguish three bandwidths (upper, middle and lower). The mean power in the mid-frequency bandwidth (0.07 – 0.14Hz) was calculated to represent a psychophysiological index of mental effort (Aasman, Mulder, & Mulder, 1987; Mulder, 1979a; Mulder, 1979b; Vicente, Thornton, & Moray, 1987).

A range of subjective measures was employed during the study. A post-test and pre-test questionnaire was administered that included the University of Wales Mood Adjective Scale (UMACL) (Matthews, Jones, & Chamberlain, 1990), the Delft Effort scale (Zijlstra & van Doorn, 1985) and the raw version of the NASA-Task Load Index (RTLX) (Byers, Bittner, & Hill, 1989; Hart & Staveland, 1988). In addition, subjects were asked to provide verbal ratings of fatigue symptoms in the course of simulated journey. These rating constituted a modified version of the fatigue symptoms checklist (Matthews & Desmond, 1998). The symptom checklist had been reduced to 16 items (constituting 4 symptoms in each of 4 categories of Visual Fatigue, Motivation, Muscular Fatigue and Malaise) where subjects had to respond "none," "low," "medium," or "high."

6.3.7 Procedure

The subjects arrived at the Institute for the initial session and received a set of standard instructions that provided an overview of the study. The subjects were fitted with ECG electrodes and asked to complete a ten-minute familiarisation journey using the simulator. Baseline data were recorded during the latter portion of this familiarisation journey.

Subjects received a set of standard instructions for each test session (control, DD, CD). These instructions included a description of the subjective data protocol, i.e. subjects instructed that they would be prompted to provide fatigue ratings at the "Services" sign. In addition, the subjects were informed about the financial rewards and penalties included in the experimental protocol. Subjects were told that they would receive a £10 minimum payment per session, but subsequent payment was contingent on the duration of time spent "driving" in the simulator. Specifically, the subjects would receive £2.50 for every 15 minute period spent performing the simulated drive. However, if they left the road with all four wheels at any time, the journey would be terminated and the subjects would forfeit all payment except the £10 minimum.

When a display was present, the standard instructions included a description of each display including graphical examples of full range of warning information, that the warning was based on their driving performance, and the degree of impairment indicated by the warning feedback (i.e. amber warning = minimum impairment of 40 per cent comparative to baseline, red warning = minimum impairment of 70 per cent relative to baseline).

Prior to beginning their journey, each subject was asked to complete a pre-test version of questionnaire index. Each subjects was instructed to drive until "they felt unable to continue at an acceptable level of performance." Subjects were provided with an opportunity to stop the session voluntarily by the appearance of a "Services" sign that appeared every 25km. When the subjects passed this sign, regardless of whether they wished to withdraw from the study, they were asked to provide subjective ratings via the fatigue symptoms checklist. If the subject continued for over 120 minutes, the session was terminated by the experimenter. Once the journey was over, the subjects completed post-test versions of questionnaire index.

This sequence of familiarisation – instructions – pre-test scales - test session – post-test scales was followed in all three experimental sessions.

6.4 Results

The data collected during the experiment were principally subjected to ANOVA procedures for the purposes of statistical analyses. Most analyses conform to a 3 x 5 design, where experimental CONDITION (Control, CD, DD) and Time-On-Task (TOT) act as main effects. Naturally, the inclusion of TOT was problematic, given that simulated journeys could be terminated at the subjects' discretion and therefore, journey duration was variable. For the purposes of analyses, it was decided to divide each journey into five proportionate periods, hence the 3 x 5 ANOVA design.

6.4.1 Total Journey Duration

The total duration of the simulated journey was analysed via a 3-way ANOVA between each of three experimental conditions. This analysis revealed no significant differences in journey duration between the Control condition (M = 97.7min) compared to those journeys when either the discrete (M = 95.7min) or the continuous display (M = 97.8min) were available.

A correlational analysis indicated that total journey duration showed only modest and insignificant levels of association with the frequency of amber and red warnings. These data are shown below in Table 9, comparative data from the Control condition are also included.

CONDITION	Amber warnings	Red warnings
CONTROL	-0.10	-0.29
DISCRETE DISPLAY	-0.32	-0.33
CONTINUOUS DISPLAY	-0.22	-0.32

Table 9. Correlation coefficients between warnings and total journey duration for all subjects (N=18). NB: None of the coefficients reached statistical significance.

6.4.2 Warning criterion and frequency of warnings

The warnings presented by the interface were based on two measures, *LPcrit* and *STVEL* (see methodology for definition). Both variables were subjected to 3 x 5 ANOVA analyses (CONDITION x TOT).

The analysis of *Lpcrit* revealed a significant effect for TOT [$F(4,68)=31.1$, $p < 0.01$] and an effect of marginal significant for CONDITION [$F(2,34)=2.9$, $p = 0.07$]. Post-hoc analyses revealed that *Lpcrit* was significantly lower during periods 1 and 2 with respect to all consecutive periods. In addition, *Lpcrit* was significantly higher during period 5 relative to periods 3 and 4. Post-hoc analyses of the CONDITION effect indicated that *Lpcrit* was significantly higher during the Control condition relative to either display condition.

The analysis of *STVEL* revealed a significant main effect due to CONDITION [$F(2,34)=4.2$, $p < 0.05$] and an interaction between both main effects [$F(8,136)=2.1$, $p < 0.05$]. The significant main effect revealed that *STVEL* was higher during the Control condition relative to either of the two display conditions. Post-hoc analyses of the interaction effect indicated that *STVEL* was significantly reduced during the DD condition during only periods 3-5. In the case of the CD condition, *STVEL* was significantly reduced during periods 2-5 relative to the Control condition. In addition, *STVEL* was only sensitive to TOT during the Control condition (i.e. *STVEL* was higher during periods 3-5 relative to 1-2). There was no TOT effect apparent during either the DD or CD conditions. This interaction effect is illustrated in Figure 31.

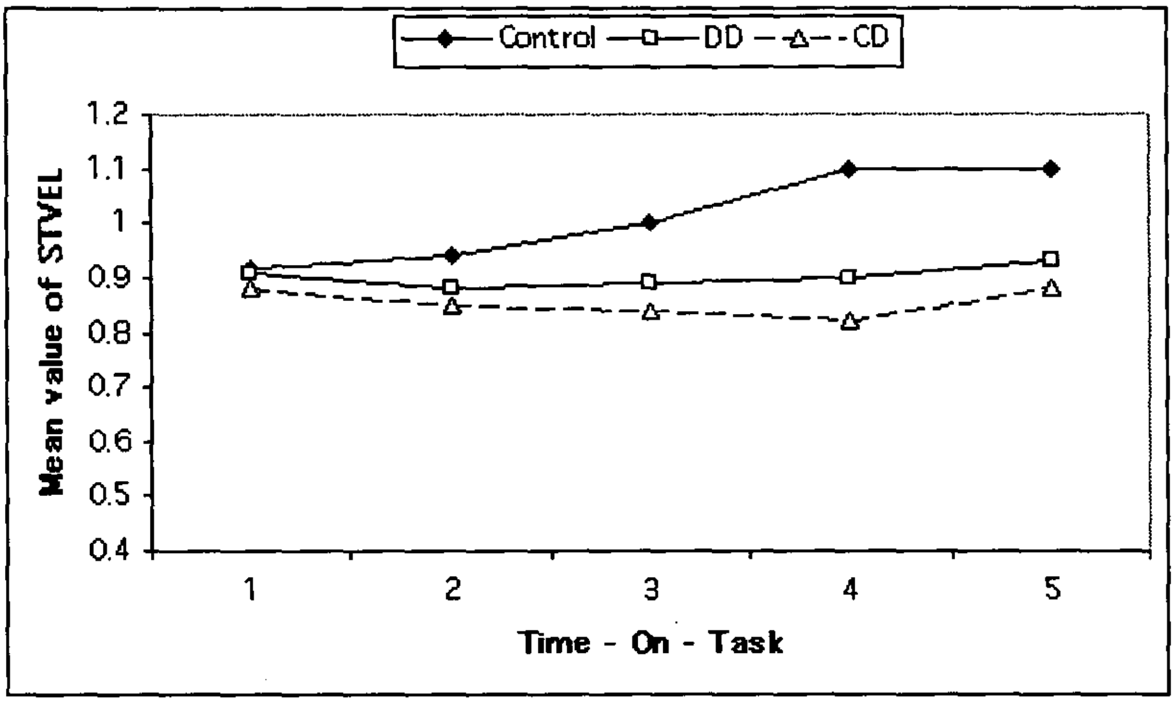


Figure 31. Mean values of *STVEL* during all 3 experimental conditions across TOT (N=18).

It was apparent from the analyses that the *Lpcrit* was the criterion measure most sensitive to TOT. Therefore, it was not surprising that the majority of warnings presented to the subjects were triggered by *Lpcrit* rather than *STVEL*, i.e. fraction of alarms due to *Lpcrit* averaged 0.98 in both display conditions, whereas the equivalent for *STVEL* were 0.6 and 0.5 for the DD and CD conditions respectively.

The warning displays in both DD and CD conditions were divided into three areas, i.e. green, amber and red. The amber zone indicated that either *Lpcrit* or *STVEL* had fallen by 40-69 per cent relative to baseline values, whereas the red zone was indicative of a primary task decrement of at least 70 per cent. The time each subject spent in either the amber or the red zones was quantified as a proportionate value and subjected to a 3 x 5 ANOVA. The analysis of time spent in the amber zone revealed significant main effects due to CONDITION [$F(2,34)=6.4$, $p < 0.01$] and TOT [$F(4,68)=7.9$, $p < 0.01$]. Post-hoc testing revealed that subjects spent proportionately less time in the amber zone in the presence of either display relative to the Control condition. In addition, it was found that subjects spent less time in the amber zone during periods 1 and 2 relative to consecutive periods.

The same analysis was performed on proportionate time spent in the red zone. In this case, there was a significant main effect due to TOT [$F(4,68)=5.1$, $p < 0.01$] and a marginal effect due to CONDITION [$F(2,34)=2.9$, $p = 0.07$]. A post-hoc analysis of the TOT effect revealed an identical trend to the one described earlier with respect to time in the amber zone. Post-hoc testing of the CONDITION effect revealed that subjects spent significantly less time in the red zone when using the CD interface ($p < 0.05$) compared to the Control condition. However, this effect was less marked when comparing the same data during the DD condition ($p = 0.09$).

Both main effects due to CONDITION for time spent in amber and red zones are shown below in

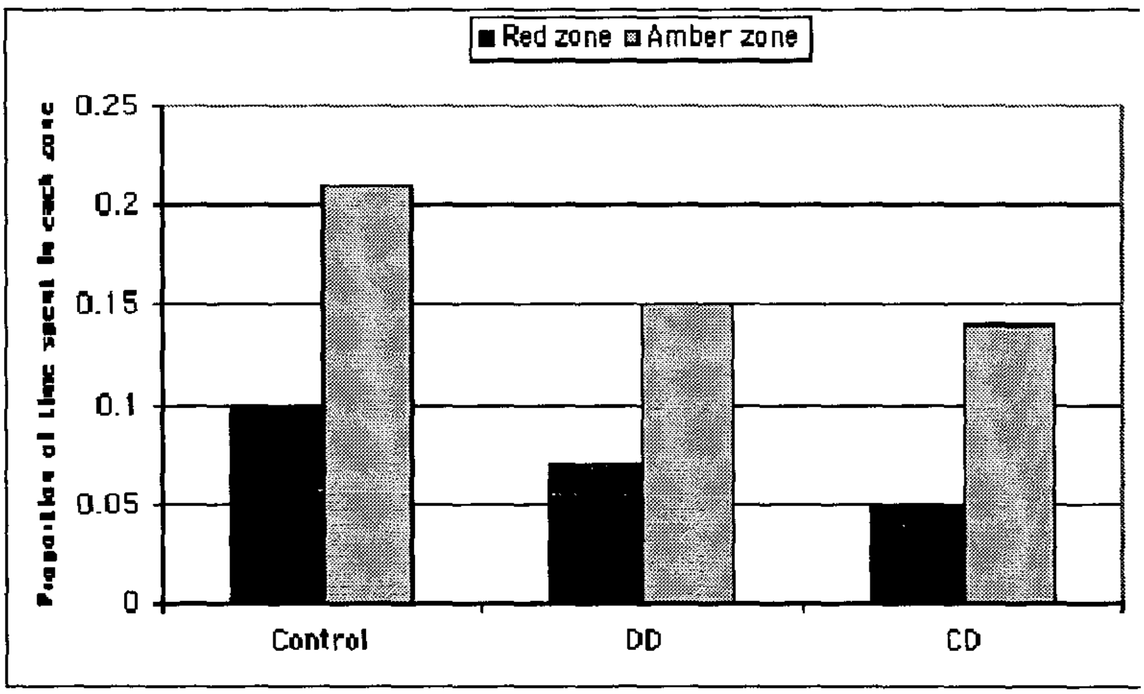


Figure 32. Mean proportion of time spent in the amber and red zones across all three experimental conditions (N=18).

6.4.3 Primary task performance

Several other aspects of vehicular control were measured and subjected to statistical analyses.

The accuracy of subjects' lateral control was indexed via two related measures: (a) frequency of lane crossings per minute, i.e. when either the right or left-side wheel came into contact with the lane boundary, and (b) frequency of near-crossings per minute, i.e. when either the right or left-side wheel was less than 20 cm from the lane boundary. These data were subjected to a 3 x 2 x 5 MANOVA (CONDITION x LATERAL CONTROL MEASURE x TOT). Several significant effects emerged from this analysis, there was a marginal effect for CONDITION [$F(2,34)=3.1$, $p = 0.06$] and significant main effects for LATERAL CONTROL [$F(1,17)=34.1$, $p < 0.01$] and TOT [$F(4,68)=15.9$, $p < 0.01$]. The main effect due to LATERAL CONTROL indicated that the near-crossing were more frequent than the occurrence of actual lane crossings. Other main effects were investigated further via a number of significant interaction effects. In the first instance, an interaction effect was apparent between CONDITION x LATERAL CONTROL [$F(2,34)=3.3$, $p = 0.05$], which indicated that there was no significant differences between the three conditions for lane crossings. However, the frequency of near-crossings was significantly lower during both display conditions relative to the Control condition. The interaction between LATERAL CONTROL and TOT [$F(4,68)=17.6$, $p < 0.01$] revealed that the frequency of lane crossings were significantly higher during period 5 relative to periods 1-2. It was also found that the frequency of near-crossings showed a significant linear trend with increasing TOT, i.e. frequency during period 1 was lower than periods 2-5, frequency during period 2 was lower than 3-5 etc. The significant main effects for near-crossings are illustrated in Figure 33.

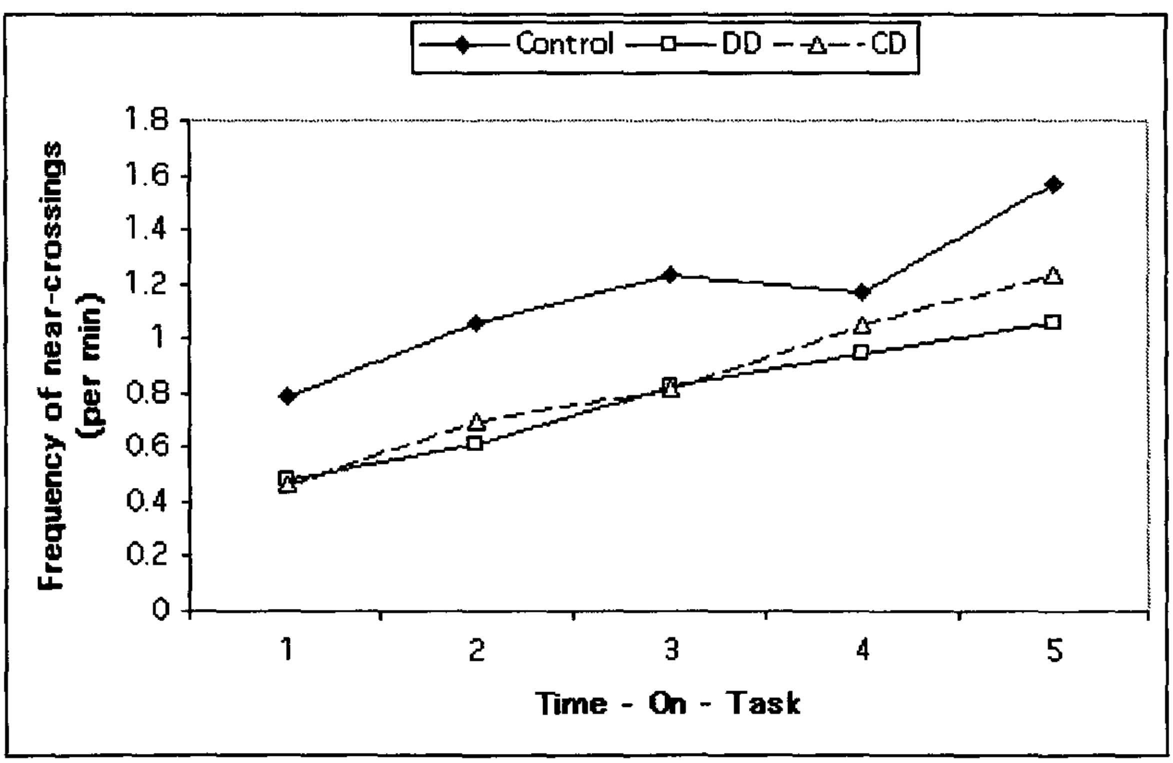


Figure 33. Mean frequency of near-crossings in three conditions across TOT (N=18).

An identical 3 x 2 x 5 MANOVA was performed to investigate the temporal dimension of both lane crossings and near-crossings, i.e. fraction of time when the 'vehicle wheels' were either in contact with, or within 20 cm of the lane boundary. This analysis revealed main effects for LATERAL CONTROL [$F(1,17)=28.1$, $p < 0.01$] and TOT [$F(4,68)=15.3$, $p < 0.01$]. These findings indicated that near-crossings were longer than actual lane crossings and that durations were lower during periods 1 and 2 relative to period 5. An interaction between CONDITION and LATERAL CONTROL [$F(2,34)=4.5$, $p < 0.05$] revealed that near-crossings were shorter in the presence of a display (this trend was not apparent for actual lane crossings).

The stability of the subjects' steering input was indexed by calculating the standard deviation of steering wheel input (steer sd). This variable was subjected to a 3 x 5 ANOVA, which revealed a significant main effect due to TOT [$F(4,68)= 7.8$, $p < 0.01$] and a significant interaction [$F(8,136)=4.1$, $p < 0.01$]. The main effect indicated the standard deviation of steering input was significantly lower during periods 1 and 2 relative to consecutive periods of TOT. Post-hoc testing of the interaction effect revealed that steer sd was significantly higher during the Control condition relative to either display conditions during periods 2-5. In addition, steer sd exhibited a linear trend over TOT during the Control condition. However, this trend was not apparent for either of the two display conditions as shown in Figure 34. The post-hoc tests indicated that steer sd was significantly higher during periods 4-5 relative to period 2 during the DD condition. There was no evidence for any such linear trend during the CD condition.

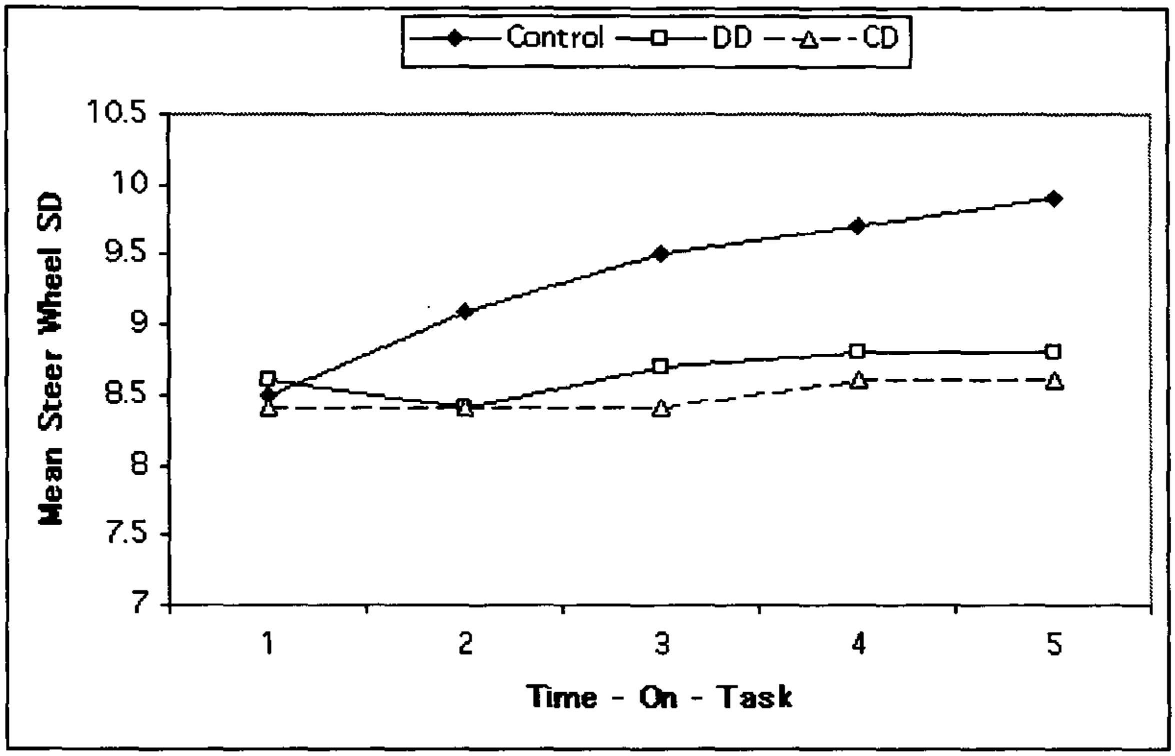


Figure 34. Mean values of steering wheel standard deviation for three experimental conditions across TOT (N=18).

The subjects' choice of mean speed was subjected to a 3 x 5 ANOVA, which revealed a main effect of only marginal significance for TOT [$F(4,68) = 2.2$, $p = 0.08$]. This effect indicated that mean speed increased with rising TOT. An identical analysis of speed variability (expressed as the standard deviation of speed) revealed no significant findings.

6.4.4 Psychophysiology

The raw ECG trace was subjected to a power spectrum analysis using the HRV extension to the Chart™ software package. Unfortunately, data from two subjects had to be discarded to measurement artefacts. Each original data file was initially sub-divided into five proportionate sections. This analysis software calculated inter-beat intervals (IBI) from the trace (for each of the five data files per subject) and the power in the mid-frequency bandwidth (0.07-0.14) of the IBI variability. These values were subjected to a natural log transform and baselined to resting values using the procedure described by Meijman (1995).

These data were subjected to a 2 x 5 ANCOVA with total journey duration acting as a changing covariate. This analysis revealed a significant main effect due to TOT [$F(4,60)= 19.0$, $p < 0.01$]. Post-hoc testing indicated that psychophysiological mental effort was higher during periods 1-2 relative to all consecutive periods.

6.4.5 Subjective data

The fatigue checklist data were summed to produce scores for each of five fatigue factors (Desmond, Matthews, & Hancock, 1997) of between 0 and 12. These summarised data were averaged into five proportionate segments of the journey and divided by the duration of each segment, i.e. to produce a fatigue score per minute. These data were subjected to 3 x 5 ANOVA. These analyses revealed only significant main effects for TOT for: Boredom [$F(4,68)= 19.5$, $p < 0.01$], Malaise [$F(4,68)= 4.4$, $p < 0.01$], Muscular fatigue [$F(4,68)= 14.4$, $p < 0.01$], Visual fatigue [$F(4,68)= 8.6$, $p < 0.01$] and the frequency of reported symptoms per minute [$F(4,68)= 13.5$, $p < 0.01$]. Post-hoc testing revealed similar linear trends with increasing TOT, e.g. a minimal number of symptoms during period 1 relative to periods 3-5, a higher number of symptoms during period 5 relative to periods 2-3.

Post-test scores from the RTLX, the UMACL and the Effort scale were subjected to a 3-way ANCOVA with total journey duration acting as a changing covariate. These analyses revealed only one significant finding for Tense Arousal [$F(2,32)= 3.3$, $p < 0.05$]. Post-hoc analysis indicated that subjects felt significantly more tense following the DD condition ($M = 18.3$) relative to either the Control or the CD condition (respective $M = 17.1$ and 17.3).

6.4.6 The decision to continue or to quit.

As stated in the Methodology section, the subjects were allowed to terminate the simulated journey at their discretion. Those subjects who did not withdraw voluntarily from the simulated task had their sessions closed by the experimenter following 120 minutes of sustained performance.

In all three conditions, approximately half the subjects were terminated by the experimenter – five of whom had their sessions closed by the experimenter in all three experimental conditions. With respect to those subjects who withdrew of their own accord, the duration of their sustained performance was very consistent between the three conditions as shown in Table 10.

	Control Condition	Discrete Display	Continuous Display
N	9	8	8
Mean duration (min)	97.6	95.7	96.8
Standard dev. (min)	27.9	25.0	27.9
Maximum duration	110	94	95
Minimum duration	36	35	26

Table 10. Descriptive statistics for subjects who withdrew from study by choice.

These data were supplemented by a discriminant function analysis (DFA), which was performed to determine which factors were associated with the decision to withdraw voluntarily from the task.

For the purposes of the DFA, the subjects were divided into two classes: a FINISH group who performed the task for the maximum duration of 120 minutes and a QUIT group who terminated the journey by their own accord.

*The following primary task variables were calculated as average values across the journey as a whole and used as independent variables in the DFA: fraction of time in yellow zone, fraction of time in red zone, frequency of lane crossing, frequency of near-lane crossing, *Lpcrit*, and *STVEL*. These primary task variables were supplemented by a number of subjective measures. These data from the fatigue symptom checklist (visual fatigue, motivation, muscular fatigue, malaise and total frequency of symptoms) were expressed as a gain rate per minute, i.e. period 5 value minus period 1 value ÷ duration of journey. The psychophysiological data representative of mental effort was also included as a gain rate per minute, i.e. mental effort at period 1 – effort at period 5 ÷ journey duration. The inclusion of the psychophysiological data within the DFA meant that two subjects were not included in these analyses due to missing ECG data.

Three DFAs were performed which employed data from the three experimental conditions. In all cases, DFAs were performed as forward stepwise procedures with F-value to enter set at 1.50 and a Tolerance value of 0.0100. The results of the DFA are shown in a standard tabular format and a note of explanation is provided for those unfamiliar with DFA. The significance of the discrimination is expressed as a global Wilks Lambda value [$\underline{\lambda}_w$], i.e. 0 = perfect discrimination. This discrimination is constituted of a model containing several variables, which are shown in the Table. The Partial Lambda represents the unique contribution to the group discrimination provided by that variable. The $\underline{F}$ ratio and associated level of probability are also provided in the Table, however since the DFA procedure capitalises on chance, these p levels should be interpreted with caution.

The DFA using data from the Control condition was terminated after six steps, resulting in a highly significant discrimination between the two groups [λ_w (6.9)=0.13, $p<0.01$]. This analysis is summarised in Table 11. According to this analysis, subjects who withdrew from the journey experienced a higher gain rate for fatigue symptoms, particularly those associated with visual and muscular fatigue. In addition, QUIT subjects showed a superior level of Lpcrit, indicating a higher accuracy of lateral control. It was apparent that the frequency of fatigue symptoms and the level of Lpcrit were the two principle factors responsible for subjects choosing to terminate the journey on a voluntary basis (i.e. see partial Lambda and F ratios in Table 11).

Variable	Partial Lambda	F-value	p-level	Mean values	
				FINISH	QUIT
Muscular fatigue	0.99	0.05	0.82	0.021	0.086
Lpcrit	0.31	25.4	<0.01	1.117	1.014
Frequency of symptoms	0.45	13.4	<0.01	0.027	0.093
Visual fatigue	0.74	3.7	0.07	0.016	0.092

Table 11. Results of the DFA analysis conducted on data from the Control condition (N=16).

The DFA based on the DD condition was terminated after two steps, resulting in a significant discrimination between the two groups [λ_w (2,13)=0.52, $p<0.05$]. This analysis revealed two factors that may have been responsible for subjects choosing to withdraw from the simulated journey (Table 12). It was apparent that subjects choose to QUIT the journey principally because of the STVEL variable, which was used as an input to the warning system. In addition, subjects who experienced the highest negative rate of gain with respect to psychophysiological effort tended to withdraw from the journey, i.e. those subjects whose effort expenditure was declining at the highest rate

Variable	Partial Lambda	F-value	p-level	Mean values	
				FINISH	QUIT
STVEL	0.71	17.9	0.04	0.759	1.073
0.1Hz mental effort	0.71	8.3	0.04	-0.171	-0.661

Table 12. Results of the DFA analysis conducted on data from the Discrete Display condition (N=16).

The DFA which was based data from the CD condition was terminated after four steps, resulting in a significant discrimination between the two groups [λ_w (4,11)=0.28, p<0.01]. This analysis is summarised in Table 13. It was apparent that three principle factors were responsible for subjects withdrawing during the CD condition. Those subjects who tended to QUIT were characterised by a higher rate of gain of those fatigue symptoms associated with declining motivation. As in the previous analysis, subjects who withdrew from the study exhibited a higher negative rate of gain associated with psychophysiological effort. It was found that QUIT subjects exhibited a lower frequency of lane crossings relative to the FINISH group. Finally, there was an indication that high levels of muscular fatigue prompted subjects to withdraw from the simulated journey by their own volition.

Variable	Partial Lambda	F-value	p-level	Mean values	
				FINISH	QUIT
Motivation	0.64	6.3	0.03	0.022	0.095
0.1Hz mental effort	0.71	4.5	0.06	-0.245	-1.031
Frequency of lane crossings	0.56	8.8	0.01	0.089	0.044
Muscular fatigue	0.85	1.9	0.19	0.025	0.114

Table 13. Results of the DFA analysis conducted on data from the Continuous Display condition (N=16).

6.4.7 Performance Efficiency Index

This analysis was based on the relationship between psychophysiological effort and performance as used in previous studies (Chapters 7 to 5) and described originally by (Meijman, 1995). For the purposes of the current analysis, the Lpcrit variable was used to represent primary task performance. The efficiency index is illustrated below in Figure 35. Values for psychophysiological effort and Lpcrit were derived for each of the five consecutive periods of the simulated journey.

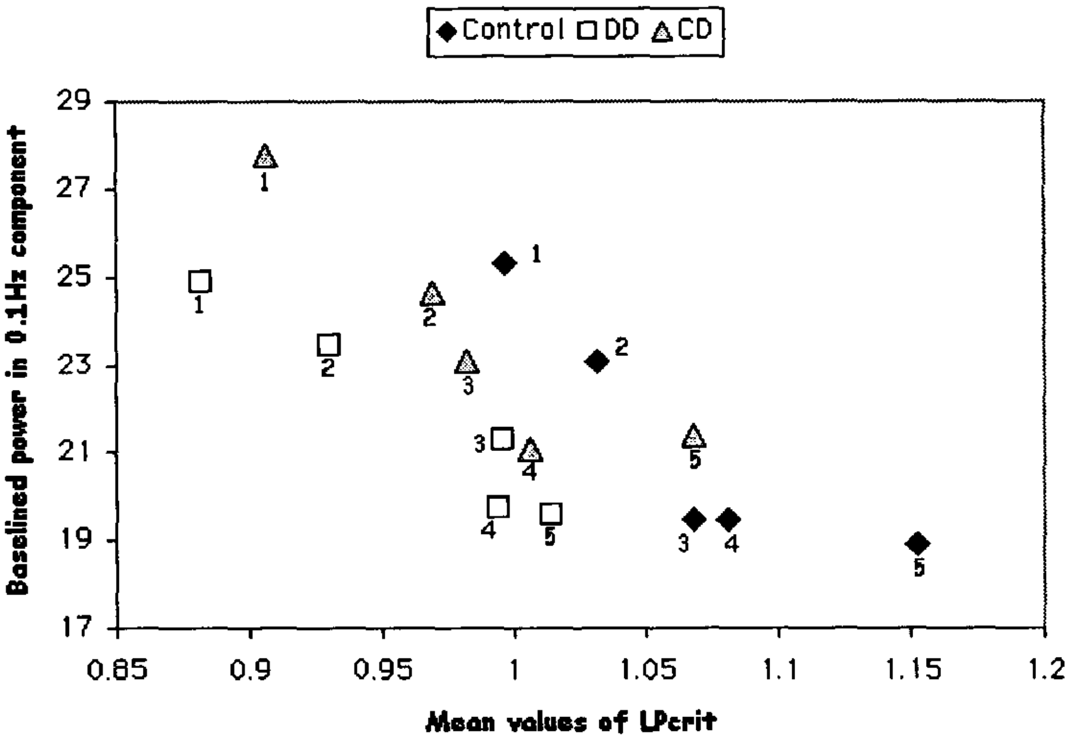


Figure 35. Performance efficiency index for all three experimental conditions. NB: numbers denote successive periods of a simulated journey, high y-values = poor performance, high x-values = increased mental effort.

It was apparent that subjects in the Control condition showed the greatest decline of performance efficiency in the course of the simulated journey. The level of mental effort fell to around 19% between periods 2 and 3 and remained at a stable, low level whilst performance declined between periods 3 and 5. In the DD condition, subjects exhibited a similar trend with respect to mental effort , i.e. falling from 25 to 19 per cent over the course of the journey. However, subjects were capable of higher overall level of performance effectiveness in the presence of the DD and therefore, sustained efficiency at a higher level throughout the journey. The presence of the CD prompted subjects to invest mental effort at a higher level during the initial period of the journey, but this increased expenditure did not translate into superior performance (relative to the DD condition). Performance efficiency during the CD condition was characterised by a slightly higher level of mental effort throughout (relative to both other conditions), however, performance was broadly at the same level as the DD condition, with the exception of periods 1 and 5.

6.4.8 The regulation of mental effort

The psychophysiological analysis revealed that the presence of performance feedback had no impact on the level of effort invested into the simulated driving task. However, it was hypothesised that feedback would influence the regulation of mental effort by increasing the salience of external performance cues. A regression analysis was conducted to test this hypothesis using psychophysiological effort as an independent variable. A number of variables represented internal costs (subjective estimates of fatigue symptoms) were entered into this analysis alongside variables representing external cues (frequency of warnings, line crossings, and near-crossings). Note that warning frequency was calculated as a "virtual" variable in the case of the Control condition.

The regression analysis of data from the Control yielded a R^2 value of 0.94 [F (7,5)= 8.6, $p < 0.05$] (three subjects rejected as outliers). The significant predictors of mental effort are shown below in Table 14. The level of association between psychophysiological effort and each variable is expressed as a semi-partial correlation coefficient. In the Control condition, mental effort significantly declined when symptoms of visual fatigue increased.

The regression analysis of DD data indicated a significant degree of association between predictor variables and mental effort. The R^2 value was 0.89 [F (7,5)= 5.9, $p < 0.05$]. The level of mental effort was significantly reduced whenever symptoms of boredom and visual fatigue were in ascendance. On the other hand, mental effort was invested when symptoms of muscular fatigue and the frequency of near line crossings increased (Table 14).

The analysis of data from the CD was insignificant, R^2 value = 0.84 [F (7,5)= 3.8, $p = 0.09$]. None of the dependent variables was found significant. However, the correlation between the frequency of near-line crossings and mental effort indicated a weak association (Table 14).

	CONTROL	DD	CD
Boredom	0.01	-0.58*	-0.19
Muscular fatigue	0.19	0.59*	0.19
Visual fatigue	-0.46*	-0.41*	0.15
Frequency of Amber warnings	-0.10	-0.29	-0.03
Frequency of Red warnings	0.27	0.26	0.02
Frequency of line crossings	0.13	0.15	-0.04
Frequency of near line crossings	-0.08	0.50*	0.30

Table 14. Semi-partial correlation coefficients between variables and psychophysiological mental effort in all three conditions (N=13).

The regression analysis revealed no direct evidence to support the hypothesis that performance feedback increased the relationship between mental effort and external cues. The level of association between warnings and mental effort was inconsistent across conditions. In general terms, there was little consensus between regression equations across the three experimental conditions.

6.5 Discussion

The study revealed that the presence of performance feedback had no effect on either the average duration of the simulated journey or the number of subjects who chose to withdraw from the task of their own accord.

Paradoxically, these subjects also exhibited a higher level of lateral control, as indicated by the *Lpcrit* data. These latter data may be used as evidence to support two hypotheses: (a) those subjects in the QUIT group withdrew earlier than FINISH subjects by definition, hence, their level of lateral control failed to decline to the same extent, i.e. Lpcrit was associated with a significant TOT effect, or (b) those subjects in the QUIT group may have been more sensitive to the level of lateral accuracy than the FINISH subjects, both in terms of external feedback and with respect to the acceptable level of performance decline. Therefore, these subjects decided to withdraw at an earlier point before lateral control was permitted to degrade to the same extent as the FINISH group.

This hypothesis was supported by the DFA on data from the DD condition (Table 12), where STVEL was one of only two significant variables that discriminated between the two subject groups. Post-hoc correlations revealed that STVEL was correlated with both the frequency of amber and red warnings (R

= 0.68 and 0.67 respectively) and so it may be assumed that frequent warnings persuaded subjects to withdraw from the journey.

The DFA summarised in Table 12 revealed that QUIT subjects exhibited a higher negative rate of gain associated with psychophysiological effort, i.e. the level of effort expenditure reduced at a higher rate across the journey duration. Post-hoc correlations indicated that mental effort was not correlated with STVEL ($\underline{R}$ = -0.11), but did reveal a substantial level of association between effort and those fatigue symptoms associated with declining motivation ($\underline{R}$ = 0.69). This association supports the hypothesis that mental effort expenditure increased in order to compensate for the accumulation of subjective fatigue. However, an increased rate of effort expenditure will exhaust finite reserves at a higher rate over the course of the simulated journey. Therefore, it is proposed that a high gain rate for subjective symptoms of fatigue accelerated the level of effort expenditure, which in turn, exhausted finite effort reserves at a higher rate per unit time. This positive feedback loop may approximate the link between effort expenditure and internal feedback of subjective discomfort in the effort regulation model.

The rank order data indicated that some subjects were more likely to respond to warning feedback in the CD condition rather than the DD condition. However, the DFA analysis from the CD condition (Table 13) did not provide any evidence that subjects withdrew from the task on the basis of warning feedback, i.e. Lpcrit, STVEL and other variables related to the feedback system were not included in the data model. The DFA analysis reinforced the role of subjective fatigue coupled to the rate of negative gain associated with psychophysiological effort (Table 13). In this case, both motivational and muscular fatigue symptoms were highly correlated with the rate of effort expenditure over TOT ($\underline{R}$ = -0.69 and 0.95 respectively). These findings are explained via the hypothesis stated in the previous paragraph. It was also apparent that QUIT subjects made fewer lane crossings per minute relative to the FINISH group. This finding was similar to the results of the DFA in the Control condition concerning the Lpcrit variable (Table 11). The two hypotheses to explain this paradoxical finding were stated earlier in this section.

The results of the DFA analyses did not reveal any conclusive differences between the reasons to quit across all three conditions. It was anticipated that variables related to the warning feedback would be implicated in the DFAs performed during the CD and DD conditions. This hypothesis only received support during the DD condition, when STVEL was a significant, discriminating variable between the two subject groups. However, this particular DFA produced the highest overall Wilks Lambda value

($\underline{\lambda}_w = 0.52$) which was indicative of a poor discrimination relative to the analyses from the Control and CD conditions ($\underline{\lambda}_w = 0.13$ and 0.25 respectively) – therefore, there are grounds to have less confidence in the results of the DFA from the DD condition.

It was apparent that the rate of effort expenditure and the rate of gain associated with subjective fatigue both played a role in the decision to quit across all three conditions. Whilst the rate of effort expenditure was only significant during the display analyses, it was highly correlated with muscular and visual symptoms of fatigue during the Control analysis (respective $\underline{R} = -0.59$ and -0.69). These findings suggested that the presence of warning feedback may have raised subjects' awareness of effort expenditure, the level of effort reserves and associated costs in terms of rising subjective fatigue. This hypothesis amounts to a claim that the presence of objective performance feedback brought the issue of effort utility to prominence, i.e. the decision whether or not to continue as assessed with respect to effort expenditure and remaining effort reserves. However, this hypothesis is highly speculative. Based on the analyses from the previous section, which postulated a cyclic link between increased effort expenditure and accelerated subjective fatigue, it is impossible to determine which of the two variables was principally responsible for the decision to quit.

The analysis of primary task performance indicated that both criterion variables (Lpcrit and STVEL) were sensitive to TOT effects during the Control condition (e.g. Figure 31). However, it would not be accurate to refer to either variable as a driver fatigue indicator. The criterion variables are simply indicative of a pattern of degraded driving behaviour, where lane weaving increased and a greater proportion of high-velocity steering corrections were made. Based on previous research (see Introduction), it would be accurate to state that this particular pattern of performance degradation has been associated with driver fatigue in the past. For the purposes of the current study, these findings demonstrated the sensitivity of simulated warning system to sustained temporal aspects of performance, which was used as a surrogate fatigue manipulation within the experimental design.

It was clear from the results that the presence of performance feedback diminished the rate of primary task degradation with respect to TOT. This finding was most apparent for those driving variables that were directly represented by the simulated warning system, e.g. Figure 31 and Figure 32. The occurrence of severe driving errors such as lane crossings (i.e. departures from the roadway) was relatively rare events that increased with TOT. However, it should be noted that the presence of warning feedback did not significantly reduce the frequency of such events. On the other hand, the presence of feedback proved a highly effective means of reducing the frequency and duration of less severe events such as near-lane crossings (Figure 33). The STVEL variable was associated with the stability of steering control input, therefore, it was not surprising that the presence of warning feedback managed to reduce the variability of steering control (Figure 34).

These findings indicated that the degradation of vehicular control due to TOT was counteracted by the provision of warning feedback. However, there was no evidence that the presence of feedback incited any adaptive safety-related response from the subjects such as speed reduction.

The analysis of psychophysiological data did not reveal any significant differences between the three experimental conditions with respect to mental effort investment. This null finding was very striking for two reasons: (a) performance feedback appeared to improve performance effectiveness without necessitating increased effort investment, and (b) the level of mental effort was found to decline with increased TOT, however this conservation of mental effort did not adversely affect performance during the latter stages of the simulated journey.

These findings suggested that mental effort may have been invested more efficiently during the display conditions. The depiction of performance efficiency (Figure 35) broadly supported this view. It was apparent that subjects invested effort at a higher level during both display conditions during periods 1-3 compared to the Control condition. It was also evident that subjects achieved a higher level of performance effectiveness during the display conditions, regardless of the level of mental effort investment. This trend was particularly marked for subjects during the DD condition in Figure 35. This pattern is indicative of improved performance efficiency.

Previous analyses (Eysenck & Calvo, 1992; Schonpflug, 1985) have stressed the quantitative basis of performance efficiency, i.e. efficiency is identified with a reduced quantity of effort and stable/improved performance effectiveness. This pattern of quantitative efficiency is apparent between the DD condition and the Control condition during periods 1-3 (Figure 35). However, during the latter periods of the journey (4-5), the level of effort investment is very similar between DD and Control conditions – but the DD group sustains performance at a much higher level. This pattern is suggestive of a qualitative form of performance efficiency. Therefore, effort is invested at the same gross level, but achieves a higher level of effectiveness via the timely introduction of procedural mental effort in response to error.

The only significant effect due the experimental condition to emerge from the analysis of the subjective data was an increased level of tense arousal in the presence of the discrete display. This finding may have indicated that increased performance efficiency (Figure 35) was only achieved at the expense of increased levels of subjective stress. The other scales included in the pre/post questionnaire were associated with null findings with respect to experimental condition manipulation. The absence of any significant differences indicated that the presence of warning feedback did not degrade alertness or cause a negative affective change. In addition, there was no evidence that either display had any effect on subjective mental workload or any sub-scale of the NASA-TLX. This effect was surprising, given the inclusion of sub-scales concerned with subjective estimates of performance quality and effort. The analysis of the fatigue symptom checklist data only revealed significant main effects due to TOT. Therefore, the presence of warning feedback neither accelerated nor reduced the rate of subjective fatigue accumulation.

There was little evidence that either type of warning interface produced superior driving performance to the other. It was anticipated that the CD version of the interface would provide greater fidelity of feedback (i.e. 9 possible zones of performance as opposed to 3 in the DD condition) but at the expense of increased visual distraction from the simulated roadway scene. There was some weak evidence to support the former hypothesis that the CD warning was the more effective at reducing the proportion of time spent in the red zone (Figure 32). However, there was no evidence for either increased subjective workload or primary task interference that may have been associated with the latter hypothesis.
The calculation of performance efficiency (Figure 35) suggested that the DD warning was associated with more efficient performance compared to the CD warning. In addition, there was some evidence from the DFA (Table 12) that subjects choose to withdraw from the journey based on DD feedback. Both effects may have been due to more concise form of warning feedback in the DD condition, i.e. subjects only received explicit feedback in the case of moderate or extreme impairment. However, this abrupt mode of discrete warning feedback may been responsible for the increase of subjective tension observed in the pre- and post-questionnaire data.

The experiment may have been improved by evaluating the influence of warning feedback under a more strenuous fatigue manipulation. For example, by extending the maximum possible duration of simulated journey from 2 hours to 4 hours and/or by running the experiment between midnight and

4AM. The current study was only capable of providing an indication of how feedback may impact on driving performance under modest fatigue regime and it is possible that any beneficial effects may evaporate under conditions that are more strenuous.

An additional flaw in the current study was the relatively small size of the subject group. For instance, in the earlier study performed by Nilsson, Nelson, & Carlson (1997), eighty subjects participated in the simulated journey. This study was designed in order to investigate idiosyncratic decision-making which constitutes the decision whether or not to continue a simulated journey and a larger subject population may have reduced inter-subject variability and permitted a more rigorous investigation of individual differences.

In addition, the level of adaptivity available to the subjects was somewhat limited by the experimental design, i.e. the subjects had a choice whether to stop or to continue. The ecological validity of this approach could be improved by offering rest breaks and financial incentives linked to journey schedules, i.e. the possibility of additional financial reward by arriving at a destination ahead of schedule being countered by financial punishment linked to accidents (hence, encouraging subjects to take more frequent rest breaks).

The current study used criterion levels of 40% and 70% to define respective, moderate and severe levels of impairment. However, this decision was arbitrary and was based on an unpublished pilot experiment. It would be interesting to vary this approach by selecting lower criterion levels of moderate and severe impairment, e.g. 15% and 40%, in order to investigate if the facilitative influence of feedback on performance may breakdown.

6.6 Conclusions

The study indicated that the provision of warning feedback had no influence on the decision to continue or to quit the simulated journey. However, it was apparent that feedback presentation prevented the degradation of driving performance due to time-on-task observed in the Control condition. It may be concluded on this basis that the precision of the external feedback loop may tend to degrade during sustained episodes of driving. This degradation may take the form of either reduced accuracy of performance appraisal or reduced fidelity of feedback associated with performance. The implication of this finding is that tired individuals may be relatively less aware of the quality of driving performance.

This conclusion is in broad agreement with the hypothesis that fatigue may degrade the process of performance appraisal (Brown, 1994; McDonald, 1987). However, there was no evidence to suggest that the appraisal of internal feedback (e.g. subjective fatigue) was influenced by the presence of performance feedback.

The data analyses indicated that feedback was capable of preventing performance degradation. This protection of primary task performance was achieved via an improved level of performance efficiency, i.e. performance was sustained at a higher level in associated with reduced or equivalent levels of mental effort expenditure compared to the Control condition. The presentation of feedback was capable of increasing performance efficiency without: (a) sustained, high expenditure of psychophysiological mental effort, (b) accelerated evolution of subjective fatigue, and (c) an increase of subjective mental workload. However, the discrete feedback condition was associated with elevated levels of subjective stress. It is concluded that feedback of performance degradation led to a timely investment of procedural mental effort, which raised and sustained performance above the zones of 'moderate' and 'severe' impairment. It is postulated that feedback influenced performance effectiveness by improving the precision of the external feedback loop, and therefore, promoting a highly efficient and exact schedule of procedural effort investment, i.e. effort was only invested when it was necessary.

It was noted that the improvement of driving performance achieved by feedback was limited to sub-critical/opaque aspects of degraded performance. For instance, feedback failed to reduce the frequency of lane crossings, which qualify as the most severe category of driver error possible in the simulated world – with the exception of an actual crash. It is postulated that feedback failed to impact on lane crossing events, because these events were obvious in each condition, i.e. subjects received noise feedback when they strayed from the lane. However, the presence of feedback did impact on the frequency and duration of near-crossing incidents, which were more frequent and less-critical in terms of 'safety'. It is logical to assume that a reduction of near-errors will lead to a decline in the frequency of actual errors. However, this logic was not supported by the experimental results and further research is necessary in this respect.

The presentation of feedback did not impact on strategic activities such as the decision to cease driving. Further analyses indicated that the decision to quit was influenced by the gain rates associated with subjective fatigue symptoms and psychophysiological mental effort. It is possible that performance feedback prompted subjects to consider their current and future levels of effort expenditure to continue driving, and that this factor may have influenced the decision to quit. However, this conclusion is highly speculative and warrants further investigation in future research.

7 A STUDY TO INVESTIGATE THE INFLUENCE OF TASK DEMANDS AND WORK SCHEDULE ON MENTAL EFFORT POLICY DURING SUSTAINED PERCEPTUAL-MOTOR ACTIVITY.

7.1 Abstract

Previous studies have investigated how individual differences, different classes of stressor and performance feedback influence the process of mental effort regulation. This study was performed to investigate how two factors (task demand, work schedule) influenced effort regulation across a sustained task. It was hypothesised that high task demands would emphasise external feedback and initiate an effort policy of investment. On the other hand, low task demands and increased time-on-task would bias regulation towards internal costs and effort conservation. Twenty-four male subjects participated in an experimental task, which constituted zero-order tracking and visual choice reaction time. The level of demand associated with the tracking task was manipulated by varying tracker target predictability, i.e. high task demands = unpredictable tracking. Subjects were divided into three groups who were exposed to three different task schedules: Continuous (CONT) subjects performed for 49 minutes under conditions of either high or low demands, Intermittent (INT) subjects were exposed to an alternating schedule of low/high demands across the 49 minute task period, whereas the Rest schedule group (REST) were allowed seven minutes of non-activity following every seven minutes of performance. Data were collected from several sources: tracking/RT performance, psychophysiology (0.1Hz bandwidth of heart rate variability) and subjective indices (the Dundee Stress State Questionnaire, subjective workload). The results clearly indicated the benefits of the REST schedule, e.g. increased performance efficiency, reduced level of energetical costs. In addition, high task demands increased the level of psychophysiological effort and associated energetical costs.

7.2 Introduction

Sustained performance may be described within a two-dimensional structure, where the level of task demand forms the x-axis and temporal scheduling constitutes the y-axis. The x-axis indicates the level of task demand and the level of effort investment necessary for requisite performance. The temporal schedule on the y-axis illustrates the duration over which requisite performance must be sustained. As effort is a finite resource, there is a negative association between time-on-task and the capacity for continued effort investment.

Hancock & Caird (1993) employed a similar two-dimensional model to describe mental workload. In this case, temporal factors were only concerned with the availability of time to reach task goals. Within industrial settings, the interaction between task demand and temporal factors is complicated by an expansive range of interacting factors. For example, task demand as indexed by subjective mental workload techniques may emphasise different facets of workload, e.g. physical demand, mental demand, temporal demand (Hart & Staveland, 1988). Similarly, temporal factors may vary with respect to different working time arrangements, e.g. hours worked, shiftwork, flexitime, overtime (Thierry & Meijman, 1994).

The purpose of the current study is to understand how task demands/workload and temporal factors interact in order to influence mental effort policy. It is proposed that increased task demands may provoke a broad strategy of effort investment. However, sustained effort investment may be compromised by a number of "costs" (Hockey, 1993; Hockey, 1997), some of which may impact on performance, e.g. subsidiary task failure, and others which influence energetical variables, e.g. increased stress via heightened sympathetic nervous activity, increased fatigue symptoms). If energetical costs reach high levels, the individual may be forced to invest higher levels of mental effort in order to protect performance from their influence. If costs reach critical levels, effort may be systematically reduced or conserved in order to alleviate associated costs such as stress and fatigue.

It is postulated that this cycle of investment and conservation represents an underlying cognitive-energetical trend with respect to task demand and temporal variables. Within the thesis literature review (Section 1.2), it has been proposed that mental effort is a finite resource, associated with a rate of expenditure during task performance and replenished during periods of rest (e.g. leisure time) and inactivity (e.g. sleep).

This finite effort resource is regulated with respect to external cues concerned with task performance and energetical costs originating from internal, psycho-physiological sources. Task demands will exert a primary influence on the former whereas time-on-task will influence the latter. The purpose of external cues related to performance is to provoke a policy of effort investment in response to increased error frequency or accelerated task demands. The investment of mental effort is associated with internal costs such as stress and fatigue. These costs fulfil an antagonistic function to inhibit sustained investment. The interaction between external cues and energetical costs determine the effort policy adopted by an individual, i.e. whether effort is invested or conserved.

The respective weighting of external and internal feedback will depend on various factors. Some of these considerations are intuitive. For example, the severity of an error detected from external sources will determine its importance relative to internal costs. An instance of near-catastrophic failure would automatically take priority over internal feedback if performance was safety-critical. This distinction is mirrored with respect to internal costs. According to the SREF model (Wells & Matthews, 1994), internal sources of stress may influence behaviour in a hierarchical fashion. This implicit hierarchy was recently operationalised by Matthews, et al. (1997) within the Dundee Stress State Questionaire, which is divided into three meta-factors of stress. The first factor of Engagement represents the level of interest and stimulation induced by the task at hand. Those costs associated with falling engagement include decreased energy and concentration. The distress factor represents potentially more damaging category of internal costs such as tension and negative affect. In this case, the individual may feel that their capacity to cope with task demands is being exceeded. When an individual suffers from those costs associated with reduced self-esteem and increased self-focus, the level of internal costs have reached the top of the hierarchy where self-image is being threatened. This state of stress represents the third meta-factor of worry.

It is proposed that all three categories of energetical costs require increased effort investment to compensate for their debilitating influence on task performance. In the case of decreased engagement, effort is required to prevent lapses of attention due to poor concentration. When an individual is significantly distressed, effort is invested to buttress performance against the influence of tension and reduced confidence. The primary influence of worry is to distract the individual from the task at hand via cognitive interference (Sarason, Sarason, & Pierce, 1990). In the former case, effort is invested to counteract the psychological inertia associated with sustained performance. When an individual is distressed or worried, it is proposed that an active process of task interference provoke additional effort investment. These mechanisms underlie the hypothesis that the purpose of energetical costs is to reduce the maximum duration of sustained effort investment.

It has been suggested that assessment of internal and external feedback is a perceptual process and subject to attentional bias. For example, it has been hypothesised that either external or internal sources of feedback may be emphasised at the expense of the other (Duval & Wicklund, 1972; Ingram, 1990). Furthermore, it may be argued that highly demanding tasks may engage the individual and emphasise external feedback at the expense of internal feedback (Matthews, Schwean, Campbell, Saklofske, & Mohamed, in press).

The role of feedback and effort regulation underlies the dynamic ebb and flow of performance when tasks are sustained. The current study was designed to investigate the interaction between task demands and time-on-task on effort regulation and performance.

7.3 Methodology

7.3.1 Experimental design.

The design was constructed as a three-way mixed within- and between-subjects design. Task demand was manipulated as unpredictable tracking (UP) and predictable tracking (P) to induce high and low levels of task demand respectively. The temporal aspects of the task were divided into seven, seven-minute segments, i.e. 49 minutes of performance in a full experimental session. Both tracking predictability and time-on-task (TOT) functioned as within-subjects variables. The between-subjects variable was provided by three different task schedules. During the Continuous schedule (CONT), subjects experienced a sustained level of either unpredictable or predictable tracking. The second schedule was Intermittent in the sense that task predictability was alternated on successive task sessions. The REST group were exposed to episodes of performance interpolated with 7 minute rest periods. Eight subjects participated in two conditions according to one of the three schedules. This design is shown below in Table 15.

Session	**1**	2	**3**	4	**5**	6	**7**	
TOT	**7m**	14m	**21m**	28m	**35m**	42m	**49m**	
Schedule 1 *Continuous*	**UP** **P**	UP P	**UP** **P**	UP P	**UP** **P**	UP P	**UP** **P**	N=8
Schedule 2 *Intermittent*	**UP** **P**	P UP	**UP** **P**	P UP	**UP** **P**	P UP	**UP** **P**	N=8
Schedule 3 *Rest breaks*	**UP** **P**		**UP** **P**		**UP** **P**		**UP** **P**	N=8

Table 15. Experimental Schedules and Design. TOT = time-on-task, P = predictable tracking, UP = unpredictable tracking.

The presentation order of task predictability/unpredictability was counterbalanced within each group. In addition, the experiment was performed at three different times of the day, morning (10:00), lunch (12:30) and afternoon (15:30). All testing times were counterbalanced across the subject group as a whole.

In Table 15, alternate task segments are shown in bold – this indicates those task sequences that were used for comparative purposes in the data analysis.

7.3.2 Apparatus and Experimental Task.

The software to present task material and to collect data was written in VisualBasic 5.0 and ran on Pentium PC computers under Windows95 on 14" colour monitors. The task had two components, a tracking task where the subjects were asked to track a lateral target movements using the mouse. Specifically, the mouse controlled a white 'subject' cursor moving within a darker rectangle representing the software-controlled cursor. If the subject cursor made contact with the computer cursor, a tone sounded and the colour of the subject cursor turned from white to red. All subjects controlled the mouse with their right hand. A depiction of the computer screen is shown in Figure 14.

The movement of the computer cursor was modelled on 16 sinusoidal waveforms, which could be manipulated with respect to frequency, amplitude and phase angle. The manipulation of frequency controlled the speed at which the cursor moved from left to right. The amplitude associated with the waveform set the maximum extent of lateral movement. The manipulation of phase angle was associated with the predictability of the computer cursor. If phase was set to zero, the movement of the cursor was perfectly predictable. This manipulation in combination with a modest amplitude and a frequency of 1 was used to induce predictable tracking. For the unpredictable condition, the phase angle was set to 12 (i.e. the 16 waves are pulled out of synchronisation with one another by a factor of 12, therefore creating erratic patterns of lateral movement to the left and to the right). In addition, the amplitude was set to maximum (i.e. the target had the fullest extent of lateral movement available) and frequency was set to 2 during unpredictable tracking, i.e. target movement was slightly faster during unpredictable tracking.

The subjects were also asked to perform a cognitive, vigilance task (Warm & Dember, 1998) in conjunction with tracking activity. The targets for this task appeared in the upper portion of the screen above the tracking task (Figure 14). The cognitive, vigilance task took the form of a letter identification activity. The subjects were presented with a Landolt C, which was shown in four different orientations (north, west, east, south). The subjects were instructed to recognise the easterly orientation as a target and all other orientations as non-targets (Figure 15).

The subjects were instructed to press the 'z' key with their left hand to indicate the presence of a target and to press the 'x' key to indicate a non-target. The vigilance stimuli were presented on a pseudo-random basis at approximately 9 events per minute and 20% of stimuli were targets. These stimuli were localised across 12 potential locations, six locations on a semi-circle slightly above the tracking cursor(s) and six on a second semi-circle closer to the top of the screen. The relationship between stimuli and stimuli localisation was randomised. These task characteristics for the RT stimuli was fixed regardless of tracking condition. A sampling rate of 1000Hz was used for data capture.

An eight-channel analogue-to-digital converter (MacLab™) was used to collect electrocardiogram data. This apparatus was connected to bioamplifiers and Chart™ software running on a Macintosh Powerbook™.

7.3.3 Subjects

24 male subjects were recruited from the local population to participate in the study. Several criteria were employed in order to guide subject recruitment, these were as follows:

- subjects should not be taking permanent medication
- subjects should be right-handed
- subjects should not be shiftworkers or work on permanent night-duty
- subjects should have not suffered from any wrist injury or RSI-related complaints
- subjects should have 20/20 vision or wear spectacles

In addition, the subjects were instructed to prepare for each experimental session as follows:

- to avoid alcohol on the night previous to an experimental session
- not to consume coffee or tea for 2 hours before the trial
- not to consume a large meal during the 2 hours before the trial

The subjects were allocated to one of the three schedule groups (CONT, INT or REST) based on age. It was felt that it was important to match the age distributions of the three subject groups as closely as possible. The age distributions for the three groups are shown below in Table 16.

Subject group	Continuous (CONT)	Intermittent (INT)	Rest breaks (REST)
Mean age	32.75	32.62	32.88
Standard deviation	7.96	7.05	9.58
Maximum	44	40	47
Minimum	22	22	20

Table 16. Age distributions of the three subject groups.

Subjects were paid for their participation at an approximate rate of £10 per hour (i.e. subjects received £35 pounds for participation in both task sessions). In addition, subjects were instructed that those two individuals from each group of eight who achieved the highest level of performance e would receive a gift voucher to the value of ten pounds.

7.3.4 Experimental Measures.

The experimental measures used during the experiment were drawn from three distinct categories of: primary task performance, psychophysiology and subjective measures.

Primary task measures included data to index both tracking and RT performance. With respect to the former, these included: Root Mean Square error (RMS), frequency of collisions and level of mouse input. For the RT task, the level of perceptual sensitivity (A') was calculated (i.e. a nonparametric version of d' (Parasuraman, 1986)) and mean reaction time (in ms) to targets and non-targets.

Psychophysiological measures constituted the electrocardiogram trace (ECG) recorded at 1000Hz. The subsequent inter-beat interval (IBI) was subjected to power spectrum analysis in order to distinguish three bandwidths (upper, middle and lower). The mean power in the mid-frequency bandwidth (0.07 – 0.14Hz) was calculated to represent a psychophysiological index of mental effort (Aasman, Mulder, & Mulder, 1987; Mulder, 1979a; Mulder, 1979b; Vicente, Thornton, & Moray, 1987).

Subjective measures were administered following sessions 1, 3, 5 and 7 (see Table 15). These measures included the full Dundee Stress State Questionnaire (Matthews, et al., 1997), the raw version of the NASA-Task Load Index (Byers, Bittner, & Hill, 1989; Hart & Staveland, 1988), a fatigue symptom checklist (Desmond, Matthews, & Hancock, 1997) and a bipolar Effort scale (Zijlstra & van Doorn, 1985).

7.3.5 Experimental Protocol

The experimental protocol differs slightly between each of the two sessions performed by the subject. On most occasions, subjects were run through the experimental protocol in pairs sitting at identical workstations in adjacent rooms. It should be made clear that subjects could neither see nor hear one another during task performance.

On arrival at the Institute , subjects received a briefing document providing broad details of the experiment and their participation. Subjects were then tested for near-vision acuity to confirm 20/20 vision and fitted with three disposable ECG electrodes. After this phase, subjects completed a short demographic questionnaire and received their training instructions. All subjects received a sequence of training sessions prior to task performance, this sequence was accompanied by a short document and was constructed as follows: (a) 3 minutes of RT task performance only (subjects did not move the mouse and concentrated on learning stimulus-response mapping for RT performance), (b) 3 minutes of predictable tracking only (instructed to ignore RT stimuli) and (c) 3 minutes of unpredictable tracking only (instructed to ignore RT stimuli). After (c), the subjects were instructed to close their eyes and to relax for 4 minutes whilst resting values of ECG activity were recorded. Following this break, subjects completed two consecutive training sessions: (d) performance of both tracking and RT task during predictable conditions for 4 minutes, and (e) performance of both tasks for 4 minutes during unpredictable task conditions.

Following the completion of training, subjects were presented with a list of instructions, which contained full details of their schedule for that particular session. Before the experimental session began, subjects were permitted a four minute familiarisation session with either predictable or unpredictable tracking, which was determined by the characteristics of their initial task session. Following the familiarisation session, subjects performed the initial seven minute task and completed the questionnaire index. The subjects were encouraged to complete the questionnaire index in less than five minutes. On completion of the questionnaire, the CONT and INT subjects performed the next consecutive task session. The REST subjects were provided with the newspaper(s) of the day and instructed to relax. This sequence of task session – questionnaire continued for 49min as shown in Table 15.

The protocol for the second subject session differed from the initial session with respect to subject training. During the second session, the subjects were instructed to relax (to collect resting ECG values) and performed training sessions (d) and (e) along with the familiarisation task. It should be noted that each subject session was only separated by 24 hours. Following completion of the second session, subjects were debriefed and paid.

7.4 Results

The study employed a mixed design of between- and within-subject variables, which were subjected to a 3 x 2 x 4 ANOVA (GROUP x DIFFICULTY x TIME-ON-TASK). However, if significant differences were present during the initial test session due to the between-subjects factor (GROUP), this was interpreted as a generic inequality between the three groups, i.e. because all subjects had the same task exposure during the initial test session. On these occasions, data obtained during sessions 3, 5 and 7 during both conditions were baselined to the initial test session in the predictable tracking condition. Therefore:

Baselinex = x(3,5,7) – x (1, predictable)

When data were baselined, a 3 x 2 x 3 ANOVA was used to estimate significant differences (GROUP x DIFFICULTY x TIME-ON-TASK).

7.4.1 Tracking performance.

Tracking performance was indexed by four variables: root mean square error (RMS error), frequency of collisions between subject-controlled cursor and tracking target, average time (in secs) to recover from a collisions (i.e. to detect and to correct for each collision) and the number of mouse movements used to control the subject-controlled tracker.

A 2 x 4 ANOVA on RMS error revealed significant main effects for both DIFFICULTY ($F[1, 21] = 336.5$, $p<0.01$) and TIME-ON-TASK (TOT) ($F[1, 21] = 4.36$, $p<0.01$), i.e. RMS error increased during unpredictable tracking and over TOT. However, there was no evidence for any significant interaction with the GROUP variable.

The same analysis was conducted on the frequency of collisions. This ANOVA also revealed a main effect for DIFFICULTY ($F[1, 21] = 622.1$, $p<0.01$) as described in the previous paragraph. In addition, there was a significant interaction between GROUP x TOT ($F[6, 63] = 2.89$, $p<0.05$), which indicated that the CON group made more frequent collisions than either of the other two groups during sessions 5 and 7. Finally, this analysis revealed a 3-way interaction between GROUP x DIFFICULTY x TOT, which was of marginal significance ($F[6, 63] = 1.99$, $p<0.09$), but is worthy of note. Post-hoc testing revealed that the CONT group made more frequent collisions than the other two groups during sessions 5 and 7 ($p<0.05$) when tracking was unpredictable. This interaction is illustrated in Figure 36 below.

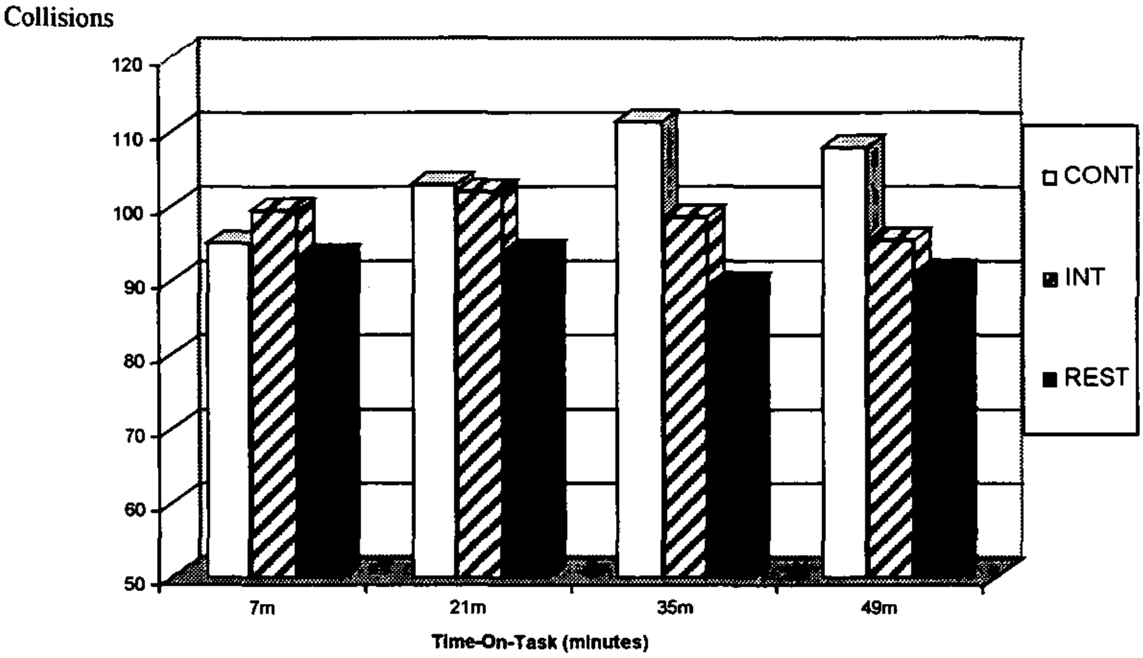


Figure 36. Mean frequency of collisions for all three experimental groups during the UNPREDICTABLE tracking condition only (n=24).

The analysis of error recovery time revealed no significant differences, whereas the ANOVA on baselined movement frequency found only a significant main effect for DIFFICULTY (i.e. higher movement during unpredictable tracking) and a marginal interaction between GROUP x DIFFICULTY ($F[2,21] = 2.7$, $p = 0.07$). The latter effect indicated that subjects in the CONT group moved the mouse more frequently during the unpredictable tracking task than the other two groups, regardless of TOT.

7.4.2 Choice reaction time.

Data from the choice reaction time task was quantified in terms of accuracy and speed of response. The former was represented by P(A), which is the non-parametric equivalent of *d'* representative of response sensitivity in signal detection theory (Parasuraman, 1986). The speed of response in terms of both targets and non-targets was measured in milliseconds.

It was found that significant differences existed between the three groups with respect to both accuracy and speed of response to visual targets. Therefore, all RT data were baselined to the initial session of the predictable tracking task for each individual subject.

The analysis of P(A) revealed a significant main effect due to DIFFICULTY ($F[1,21]=4.45$, $p = 0.05$). It was apparent that P(A) fell by approximately 0.02 during the unpredictable tracking task compared to a substantive decline of 0.04 during the predictable tracking task. Therefore, sensitivity to visual targets appeared lower during predictable tracking comparative to unpredictable tracking. In addition, this analysis revealed a 3-way interaction effect ($F[6, 63] = 7.1$, $p<0.01$), which is illustrated in Figure 37.

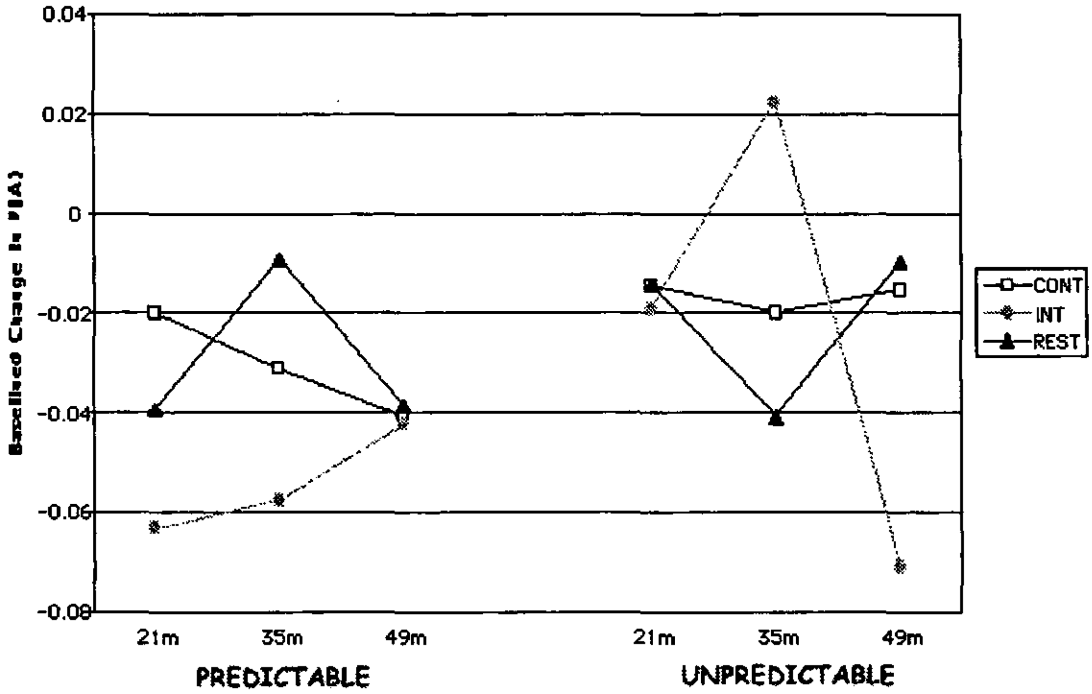


Figure 37. Baselined values of P(A) for all three groups of subjects across both conditions (N=24). Note: negative values = declining sensitivity.

Post-hoc testing revealed that baselined P(A) was significantly higher for the CONT group compared to INT group after 21 minutes during the predictable tracking task. In addition, P(A) was superior for the REST group relative to the INT group following 35 minutes of predictable tracking. During the unpredictable tracking task, the INT group exhibit superior sensitivity following 35 minutes relative to other two groups. However, the sensitivity of the INT group also showed a significant decline after 49 minutes relative to both the CONT and the REST groups.

An identical ANOVA was conducted on mean reaction time latency to target items. This analysis revealed a significant main effect for DIFFICULTY, i.e. subjects tended to respond more rapidly during the unpredictable tracking task relative to the predictable tracking task. In addition, a 3-way interaction of marginal significance ($F[6, 63] = 1.78$, $p=0.1$) was apparent. This interaction is illustrated in Figure 38. Post-hoc testing revealed that subjects in the REST group took longer to respond to targets during predictable tracking relative to the other two groups ($p < 0.05$) after 35 minutes of performance. With respect to the unpredictable tracking condition, both the INT and the REST group exhibited shorter response latencies relative to the CONT group after 35 minutes of performance.

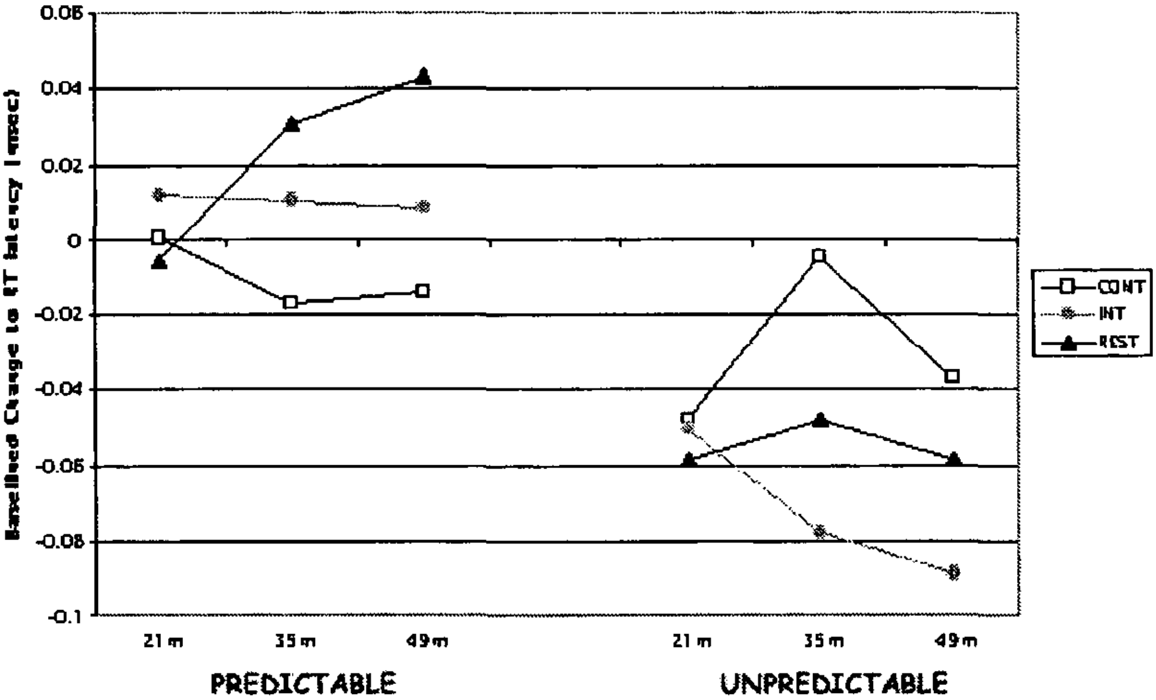


Figure 38. Baseline change in target response latency for all three groups of subjects in both tracking conditions (N=24). Note: positive values = increased latency.

An identical analysis of response latency was conducted with respect to non-target identification. No significant differences were apparent.

7.4.3 Psychophysiology

The raw ECG data were subjected to analysis using the Heart Rate Variability extension software with Chart 1.5.6. This analysis performed artefact detection/correction and power spectrum analysis in order to isolate the mid-frequency component of the heart rate variability signal (0.1Hz component). These data were obtained for a rest period prior to performance as well as experimental task sessions. The mean power from the mid-frequency bandwidth was subjected to natural log transformation and a baselining procedure, where the 0.1Hz component during performance was subtracted from the rest value, i.e. positive change is synonymous with increased mental effort (Meijman, 1995). These transformed data were subject to a 3 x 2 x 4 ANOVA (GROUP x DIFFICULTY x TOT).
This analysis revealed only significant main effects for both DIFFICULTY ($F[2,21] = 7.1$, $p<0.05$) and TOT ($F[3,63] = 7.3$, $p<0.01$). Post-hoc testing revealed a higher level of mental effort during unpredictable tracking relative to the predictable task. In addition, mental effort was significantly higher during the initial 7 minute period relative to the following three task intervals during unpredictable tracking, and during the 21 minute period relative to the following two task intervals in the predictable condition; both effects are illustrated in Figure 39.

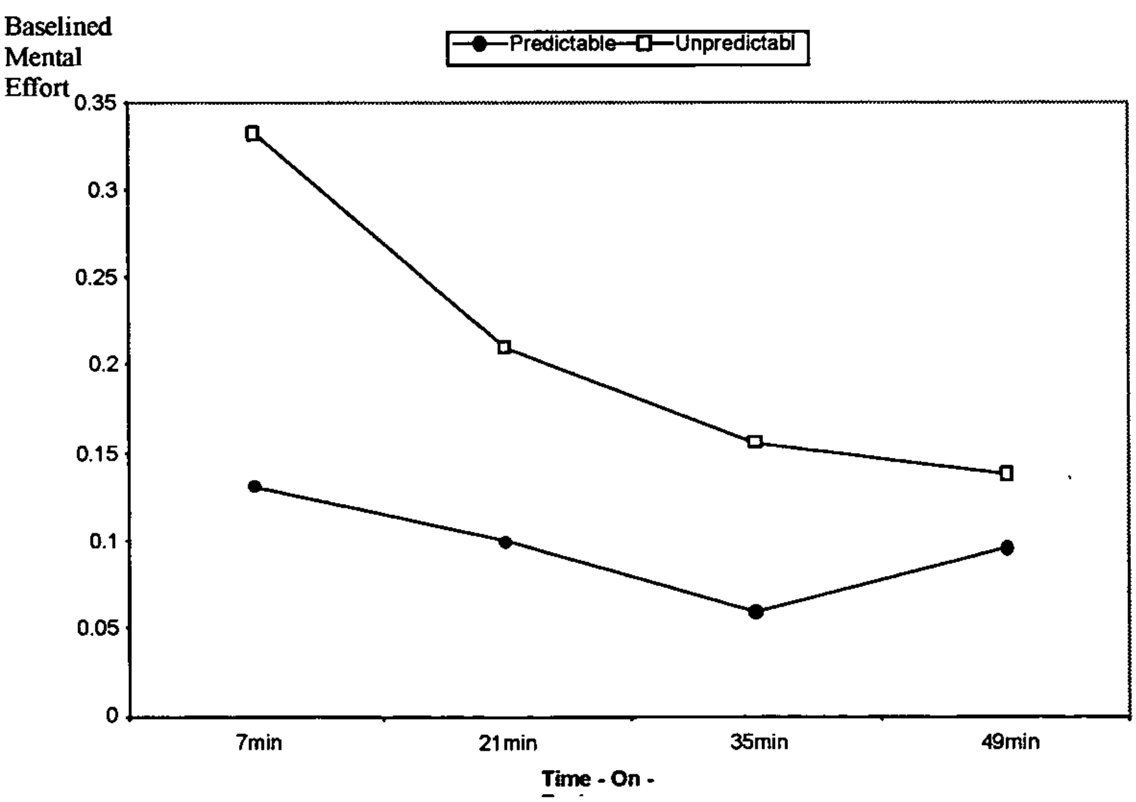


Figure 39. Mean power in 0.1Hz bandwidth (following log-transform and baselining to rest period) for both tracking conditions across time-on-task (N=24).

7.4.4 Subjective data.

The subjective data was analysed in terms of six basic groups, these were: goal-setting, the fatigue symptom checklist devised by Desmond, Matthews, & Hancock (1997), a raw version of the NASA-Task Load Index (RTLX) (Byers, Bittner, & Hill, 1989; Hart & Staveland, 1988) and the Dundee Stress State Questionnaire described by (Matthews, et al., 1997), which may be sub-divided into three categories of Engagement, Distress and Worry.

The one exception to this categorisation was the mental effort scale devised by (Zijlstra & van Doorn, 1985). These data were analysed via a 3 x 2 x 3 ANOVA (GROUP x DIFFICULTY x TOT) using baselined data, i.e. data converted to a change score based on deviation from baseline value collected following the initial task session. This analysis revealed main effects for GROUP and DIFFICULTY variables. The REST group rated their mental effort as significantly lower ($\underline{M}$ = 10) than either the CONT or INT groups ($\underline{M}$ = 29.3 and 30 respectively). In addition, mental effort was rated as significantly higher during the unpredictable tracking task relative to the predictable task ($\underline{p}$ < 0.01).

Goal-setting and self-rated performance

Goal-setting behaviour was measured in terms of three variables: (a) aspired goal level, i.e. the standard of performance subjects set for themselves prior to task session, (b) achieved goal level, i.e. the standard of performance subjects felt they had achieved following the task session, and (c) subjective goal-performance discrepancy, i.e. aspired goal standard prior to session minus achieved goal standard. All goal-related data were assessed on a four-point scale with reference to five qualitative aspects of performance: consistency, speed of response (to RT targets), efficiency, accuracy and quality. The raw data from these scales were analysed via a 3 x 2 x 3 ANOVA (GROUP x DIFFICULTY x TOT) and the results are shown in

Table 17 below.

Table 17. ANOVA results for pre-test ratings of aspired goal levels (n=24).

Variable	GROUP A	DIFFICULTY B	TOT C	INTERACTIONS
Consistency	$F(2,21)=6.6$ $p < 0.01$	$F(1,21)=35.3$ $p < 0.01$	$F(2,42)=3.3$ $p < 0.05$	A x B $F(2,21)= 2.9$ $p = 0.06$
Speed		$F(1,21)=28$ $p < 0.01$	$F(2,42)=14.7$ $p < 0.01$	A x B $F(2,21)=3.6$ $p < 0.05$ A x C $F(4,42)=3.9$ $p < 0.05$
Efficiency	$F(2,21)=3.7$ $p < 0.05$	$F(1,21)=21.9$ $p < 0.01$	$F(2,42)=5.8$ $p < 0.05$	
Accuracy		$F(1,21)=27.5$ $p < 0.01$	$F(2,42)=8.77$ $p < 0.01$	
Quality	$F(2,21)=4.1$ $p < 0.05$	$F(1,21)=29.8$ $p < 0.01$	$F(2,42)=6.95$ $p < 0.01$	

These analyses revealed that the REST group set goals significantly higher with respect to the consistency, efficiency and quality of performance. The mean values for pre-test ratings of aspired goal levels are shown below in

GROUP	CONSISTENCY	EFFICIENCY	QUALITY
CONTROL	3.57	3.30	3.38
INT	2.81	2.71	2.86
REST	2.84	2.56	2.64

Table 18. Mean values of aspired goal levels by experimental group (n=24).

The main effects due to DIFFICULTY shown in

Table 17 indicate that subjects lowered their goal aspiration during the unpredictable tracking relative to the predictable tracking conditions (respectives $\underline{M} = 2.56$ and 3.34). In addition, main effects due to TOT indicate a declining level of goal aspiration as result of sustained performance, i.e. $\underline{M} = 3.21$ prior to task activity compared to 2.74 prior to final period of the task activity. Post-hoc testing revealed that significant interactions between GROUP x DIFFICULTY (for goals related to the consistency and speed of performance) both indicated that the REST group set these goals higher than the other two groups, but only during unpredictable tracking. Similarly, the significant interactions between GROUP x TOT for speed indicated that the REST group sustained a higher level of goal-setting following 21 minutes of performance relative to the other two groups.

A series of analyses were conducted to investigate the second aspect of goal behaviour – self-assessed aspects of achieved performance, which were appraised immediately following the session. The results of these analyses are shown below in Table 19

Variable	GROUP A	DIFFICULTY B	TOT C	INTERACTIONS
Consistency	$F(2,21)=4.1$ $p < 0.05$	$F(1,21)=58.1$ $p < 0.01$	$F(2,42)=4.6$ $p < 0.05$	
Speed		$F(1,21)=32.8$ $p < 0.01$	$F(2,42)=3.4$ $p < 0.05$	
Efficiency		$F(1,21)=64.4$ $p < 0.01$	$F(2,42)=4.4$ $p < 0.05$	A x B x C $F(6,63)=2.2$ $p < 0.05$
Accuracy		$F(1,21)=45.5$ $p < 0.01$	$F(2,42)=9.5$ $p < 0.01$	
Quality	$F(2,21)=3.2$ $p < 0.05$	$F(1,21)=36.7$ $p < 0.01$	$F(2,42)=5.5$ $p < 0.01$	

Table 19. ANOVA results for post-task ratings of achieved performance (n=24).

These analyses clearly indicated that subjects were aware of declining performance standards during unpredictable tracking and with increasing time-on-task. With respect to the latter, subjects generally perceived standards to significantly decline after 21 minutes of performance. In addition, subjects in the REST group perceived both the quality and the consistency of their performance to be superior to the CONT group ($p < 0.05$). The three-way interaction with respect to efficiency indicated that the CONT group rated their efficiency as lower than the REST group after 35 minutes of performance, regardless of tracking condition. In addition, the INT group rated their efficiency as lower than the REST subjects following 49 minutes of predictable tracking performance. There was no significant difference between the CONT group and the INT group with respect to self-rated efficiency.

An identical series of analyses was conducted with respect to the goal discrepancy data (i.e. pre-test aspiration minus post-test appraisal of goal attainment). The results of these ANOVA revealed that unpredictable tracking produced significantly higher goal discrepancies with respect to consistency ($F(1,21)=6.5$, $p<0.05$) and efficiency ($F(1,21)=5.6$, $p<0.05$), relative to the predictable condition., i.e. unpredictable tracking produced a higher magnitude of goal-performance discrepancy.

Fatigue symptoms

The fatigue symptoms checklist was quantified in terms of: (a) total number of fatigue symptoms reported by subjects, and (b) as a component score between 0-12 on each of four subcomponents described by Desmond, Matthews, & Hancock (1997) (visual fatigue, malaise, muscular fatigue, boredom). In both cases, data were baselined to the initial test session and analysed via 3 x 2 x 3 ANOVA. The results of this analysis are shown in Table 20 below. The mean values for all variables are provided in the Appendices.

Variable	GROUP A	DIFFICULTY B	TOT C	INTERACTIONS
Number of symptoms			$F(2,42)=11.9$ $p < 0.01$	A x C $F(2,21)= 3.9$ $p = 0.01$
Visual fatigue			$F(2,42)=12.8$ $p < 0.01$	A x B x C $F(4,42)=3.1$ $p < 0.05$
Malaise			$F(2,42)=5.7$ $p < 0.01$	
Muscular fatigue		$F(1,21)=7.1$ $p < 0.05$	$F(2,42)=13.1$ $p < 0.01$	
Boredom			$F(2,42)=3.5$ $p < 0.05$	

Table 20. ANOVA results from analysis of fatigue symptom checklist data (n=24).

These results revealed a significant main effect for TOT for all variables associated the fatigue symptom checklist, i.e. fatigue increased over TOT. The interaction between GROUP x TOT indicated that the REST group reported significantly fewer fatigue symptoms following 21 minutes of the experimental session. The significant 3 x 2 x 3 interaction observed for the visual fatigue scale was subjected to post-hoc analysis. These results indicated that the INT group reported higher levels of visual fatigue after 21 minutes of performance relative to the REST group during predictable tracking. In addition, it was apparent that both the CONT and INT groups experienced increased visual fatigue following 21 minutes of unpredictable tracking comparative to the REST group. The only main effect due to DIFFICULTY was a significant increase of muscular fatigue during unpredictable tracking.

Engagement

The engagement factor is composed of three distinct sub-scales: (a) intrinsic motivation, i.e. related to interest in the experimental task (Desmond, 1998), (b) energetical arousal, i.e. bipolar mood scale indicative of a continuum between alertness and tiredness (Matthews, Jones, & Chamberlain, 1990), and (c) level of concentration devoted to task (Matthews, et al., 1997). All three components were baselined and subjected to 3 x 2 x 3 ANOVAs (see Table 21

Variable	GROUP A	DIFFICULTY B	TOT C	INTERACTIONS
Intrinsic motivation				
Energetic arousal		$F(1,21)=4.7$ $p < 0.05$	$F(2,42)=5.3$ $p < 0.01$	
Concentration	$F(2,21)=3.2$ $p < 0.05$		$F(2,42)=10.3$ $p < 0.01$	

Table 21. ANOVA results from analysis of Engagement subcomponents (n=24).

These results indicated reduced levels of energetic arousal during unpredictable tracking (M = 14.6 during UP task compared to 16 during P task) and a reduction of both energy and concentration with increasing time-on-task. In addition, it was apparent that levels of concentration were significantly higher for the REST group compared to the other two groups, i.e. M = 21.3 for REST group compared to respective M of 20 and 20.4 for CONT and INT groups.

Distress

This factor is composed of three distinct sub-scales as described by Matthews, et al. (1997), they are: (a) tense arousal, i.e. bipolar mood scale indicative of a continuum between relaxation and tension, (b) hedonic tone, i.e. bipolar mood scale indicative of a continuum between happiness and sadness (Matthews, Jones, & Chamberlain, 1990), and (c) level of confidence in ones ability to perform the task (Matthews, et al., 1997). All three components were baselined and subjected to 3 x 2 x 3 ANOVAs (see Table 22 below).

Variable	GROUP A	DIFFICULTY B	TOT C	INTERACTIONS
Tense arousal		$F(1,21)=10.9$ $p < 0.01$		
Hedonic tone		$F(1,21)=14.4$ $p < 0.01$	$F(2,42)=8.6$ $p < 0.01$	
Confidence		$F(1,21)=4.1$ $p < 0.05$		A x B $F(2,21)= 3.1$ $p = 0.06$ A x B x C $F(4,42)=2.6$ $p < 0.05$

Table 22. ANOVA results from analysis of Distress subcomponents (n=24).

A significant main effect for DIFFICULTY was apparent across all three Distress sub-components, which was indicative of increased tension and decreased positive affect/confidence when subjects performed the unpredictable tracking task. In addition, there was a significant main effect for hedonic tone due to TOT, i.e. a declining level of positive affect following 35 minutes relative to 21 minutes of performance. The mean values for the DIFFICULTY main effect are shown below in Table 23.

Variable	Predictable	Unpredictable
Tense Arousal	14.10	15.79
Hedonic tone	25.5	23.4
Confidence	16.1	14.9

Table 23. Mean values for Distress variables by task demand (n=24)

A number of interaction effects were observed with respect to the Confidence sub-scale. Post-hoc testing revealed the CONT group rated their level of confidence as significantly lower than participants in the other two groups during predictable tracking, but only following 21 minutes of performance.

Worry

There are five distinct sub-scales subsumed within the Worry factor[11] (Matthews, et al., 1997), these are as follows: (a) extrinsic motivation, i.e. related to the need to succeed and fear of failure, (b) frequency of task-relevant thoughts, i.e. derived from the Cognitive Interference questionnaire produced by (Sarason, Sarason, Keefe, Hayes, & Shearin, 1986), (c) frequency of task-irrelevant thoughts, i.e. also derived from (Sarason, Sarason, Keefe, Hayes, & Shearin, 1986), (d) self-focus, i.e. level of self-directed attention derived from (Fenigstein, Scheier, & Buss, 1975), and (e) self-esteem, i.e. derived from (Heatherton & Polivy, 1991). All variables were baselined and subjected to 3 x 2 x 3 ANOVAs, the results of these analyses are shown in Table 24.

Variable	GROUP A	DIFFICULTY B	TOT C	INTERACTIONS
Extrinsic motivation			$F(2,42)=3.1$ $p = 0.06$	A x B $F(2,21)= 3.2$ $p < 0.01$
Task-relevant thoughts	$F(2,21)=3.2$ $p < 0.05$	$F(1,21)=8.5$ $p < 0.01$		A x B $F(2,21)= 3.3$ $p < 0.01$
Task-irrelevant thoughts			$F(2,42)=8.6$ $p < 0.01$	A x B x C $F(4,42)=2.7$ $p < 0.05$
Self-focus				
Self-esteem				A x C $F(4,42)=3.5$ $p < 0.05$

Table 24. ANOVA results from analysis of Worry subcomponents (n=24).

These analyses revealed that increased TOT reduced subjects' levels of extrinsic motivation and increased the frequency of task irrelevant thoughts. With respect to the former sub-component, it was apparent that the CONT group expressed a significantly lower level of extrinsic motivation during both

[11] Note this characterisation of the Distress factor described by Matthews et al (1997) differs from the original with respect to the inclusion of the additional factor of extrinsic motivation derived from Desmond (1998).

predictable and unpredictable tracking conditions. In addition, the INT group expressed a higher level of extrinsic motivation during predictable tracking relative to either the CONT or the INT group.

The significant main effect due to GROUP observed for the task-relevant (TR) thoughts revealed a increased frequency of TR thoughts for the INT group relative to both other subject groups. In addition, task-relevant thoughts tended to significantly decline during predictable tracking comparative to the unpredictable tracking condition. Post-hoc testing of the significant 3 x 2 interaction confirmed the significant main effect of GROUP during the predictable tracking condition. However, it was also revealed that the INT group showed significantly higher number of TR thoughts during unpredictable tracking relative to both other groups. Mean values for the analysis of TR items are provided below in Table 25

GROUP	Group mean	Predictable	Unpredictable
CONT	22.6	16.4	22.6
INT	25.9	15.7	20.5
REST	22.3	14.9	14.6

Table 25. Mean values for frequency of task-relevant thoughts for all three experimental groups (n=24).

It was apparent that the number of task-irrelevant (TI) thoughts increased with sustained TOT. Post-hoc testing of the 3-way interaction revealed that INT experienced more TI thoughts during both predictable and unpredictable tracking conditions relative to the CONT group. In addition, the frequency of TI thoughts was significantly higher for INT group than the REST group during predictable tracking and following 21 minutes of unpredictable tracking. During predictable tracking, it was found that the number of TI thoughts were higher for the CONT group relative to the REST group following 35 minutes of performance.

A significant interaction between GROUP x TOT was found for the subjective self-esteem variable. Post-hoc testing indicated that self-esteem was maximised after 21 minutes of performance during predictable tracking for the CONT group. However, this trend was reversed during unpredictable tracking and self-esteem peaked at 49 minutes, once the task had been completed.

Subjective mental workload

The Raw Task Load IndeX (RTLX) contains seven sub-scales that represent the subjects' impression of the mental workload associated with task performance, i.e. mental demand, physical demand, temporal demands, performance, effort and frustration. These scales may be averaged to provide a global index of subjective mental workload (SMW). Each scale and the global index were baselined to the initial session involving predictable tracking and analysed via a 3 x 2 x 3 ANOVA. The results of these analyses are shown in Table 26. Mean values for the RTLX items are provided in the Appendices.

Variable	GROUP A	DIFFICULTY B	TOT C	INTERACTIONS
Mental demand	$F(2,21)=2.7$ $p < 0.05$	$F(1,21)=36.8$ $p < 0.01$	$F(2,42)=9.65$ $p < 0.01$	A x B x C $F(4,42)=2.4$ $p < 0.05$
Physical demand		$F(1,21)=11.8$ $p < 0.01$		
Temporal demand		$F(1,21)=36.4$ $p < 0.01$		
Performance		$F(1,21)=12.6$ $p < 0.01$		
Effort		$F(1,21)=24.6$ $p < 0.01$		
Frustration		$F(1,21)=12.7$ $p < 0.01$		
Subjective mental workload		$F(1,21)=28.1$ $p < 0.01$	$F(2,42)=3.6$ $p < 0.05$	

Table 26. Results of Raw Task Load Index (RTLX) analyses (N=24).

The analyses of subjective mental workload clearly illustrated that workload increased during unpredictable tracking relative to predictable tracking on all RTLX sub-scales. The mental demand sub-scale appeared particularly sensitive to the experimental manipulations. The GROUP effect for this factor indicated that mental demand was significantly lower for the REST group relative to the other two groups ($p < 0.05$). In addition, mental demand significantly increased with TOT regardless of GROUP or DIFFICULTY. The 3-way interaction was subjected to post-hoc testing. This analysis revealed significantly higher levels of mental demand reported by the CONT group relative to the INT group following 35 minutes of unpredictable tracking performance.

7.4.5 Performance Efficiency index

The efficiency of performance was operationalised via a consideration of the relationship between overt indices of performance and mental effort as indexed by psychophysiology (Meijman, 1995). For the purpose of the current study, it was desirable to investigate efficiency with respect to both tracking and reaction task performance in conjunction with mean power in the 0.1Hz bandwidth of heart rate variability.

The first efficiency index was calculated to illustrate the relationship between tracking ability and psychophysiological effort. The 0.1Hz data were expressed as a percentage change from rest values, i.e. increasing positive percentage values being indicative of increased mental effort (Meijman, 1995). A similar transformation was performed with respect to RMS error data, these values were expressed as a percentage change from values obtained during the initial predictable tracking period. In this case, increasing positive percentage values were synonymous with declining performance, i.e. increased RMS error. The relationship between tracking ability and psychophysiological effort is shown below in Figure 40 for the predictable tracking task.

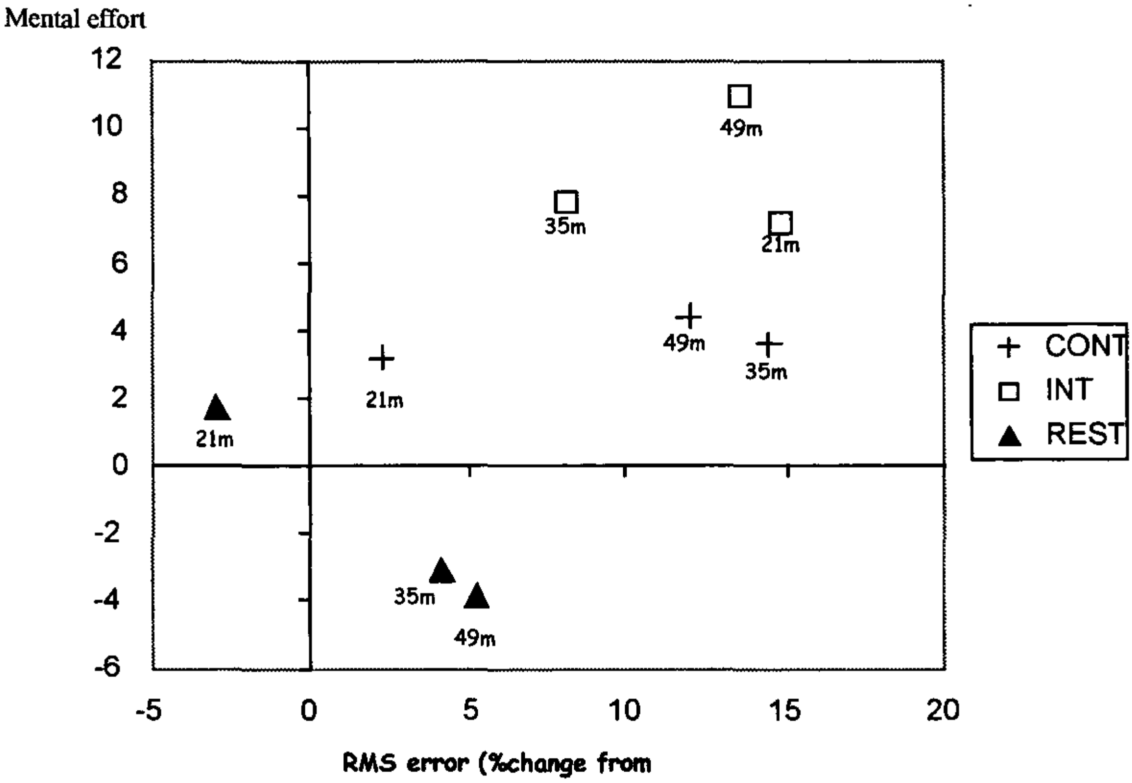


Figure 40. Performance efficiency index for tracking ability (%baseline RMS error) and psychophysiological mental effort (%baseline power in 0.1Hz bandwidth) during PREDICTABLE tracking (N=24). NB: Positive x = increased mental effort from baseline, positive y = increased tracking error from baseline.

Figure 40 illustrates a number of differences between the three experimental groups with respect to tracking efficiency. The CONT group appear to conserve mental effort (<5% from resting values) whilst tolerating comparatively high levels of RMS error. A similar level of RMS error was apparent for the INT group, but these subjects consistently invested a higher level of mental effort across each task period. By contrast, the REST group were conserving mental effort, which actually fell below resting values during the latter two task periods – however, this group were simultaneously capable of the highest overall level of tracking performance.

The same analysis was conducted on data from the unpredictable tracking task, these data are illustrated in Figure 41. In this case, the INT group invested the highest level of mental effort and obtained the lowest overall RMS error. However, the REST group only increased effort by approximately 4% and were capable of almost equivalent performance. The CONT group did not invest high levels of effort and sustained the highest level of RMS error through the task periods.

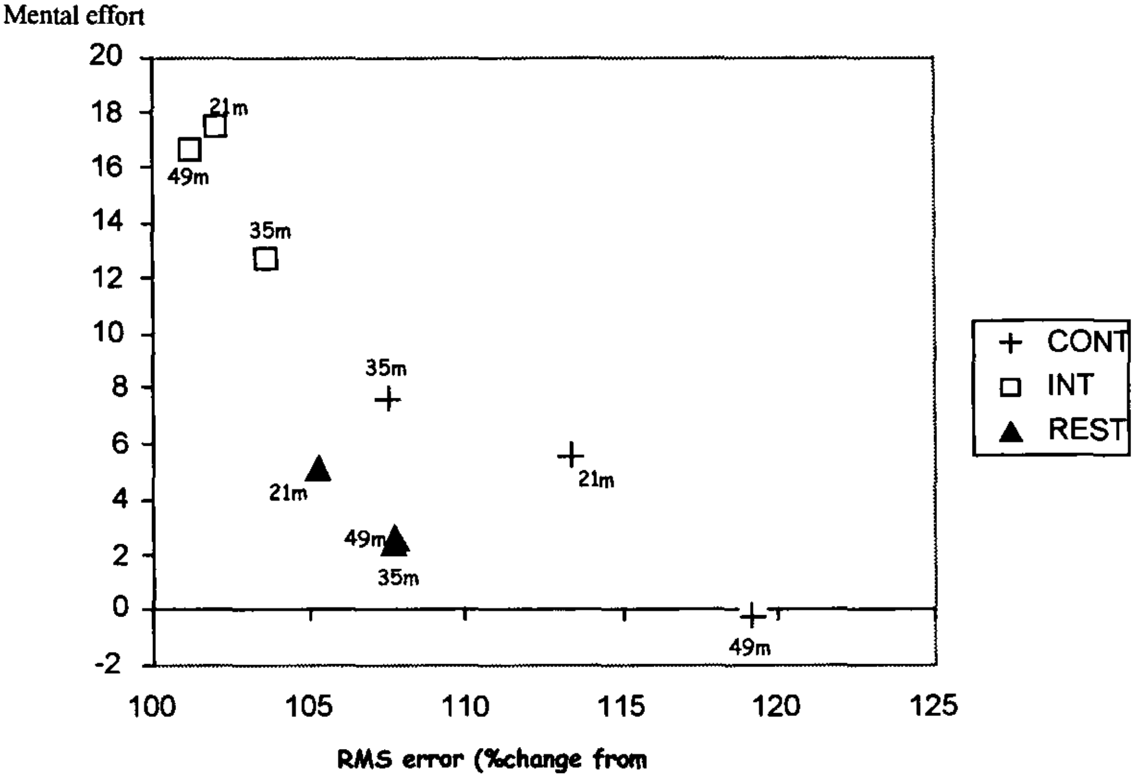


Figure 41. Performance efficiency index for tracking ability (%baseline RMS error) and psychophysiological mental effort (%baseline power in 0.1Hz bandwidth) during UNPREDICTABLE tracking (N=24).

An identical analysis was conducted to investigate the relationship between mental effort and performance on the reaction time task. In this instance, obtaining an adequate index of performance was complicated by the fact that both response latency and target accuracy are representative of overt output. The former was baselined to values obtained during the initial period of the predictable tracking task, and expressed as a percentage change value. The latter was quantified as the percentage of incorrect target responses, including both false alarms and misses. These data are shown for the predictable tracking condition in Figure 42. This analysis and representation is based on Meijman, (1995).

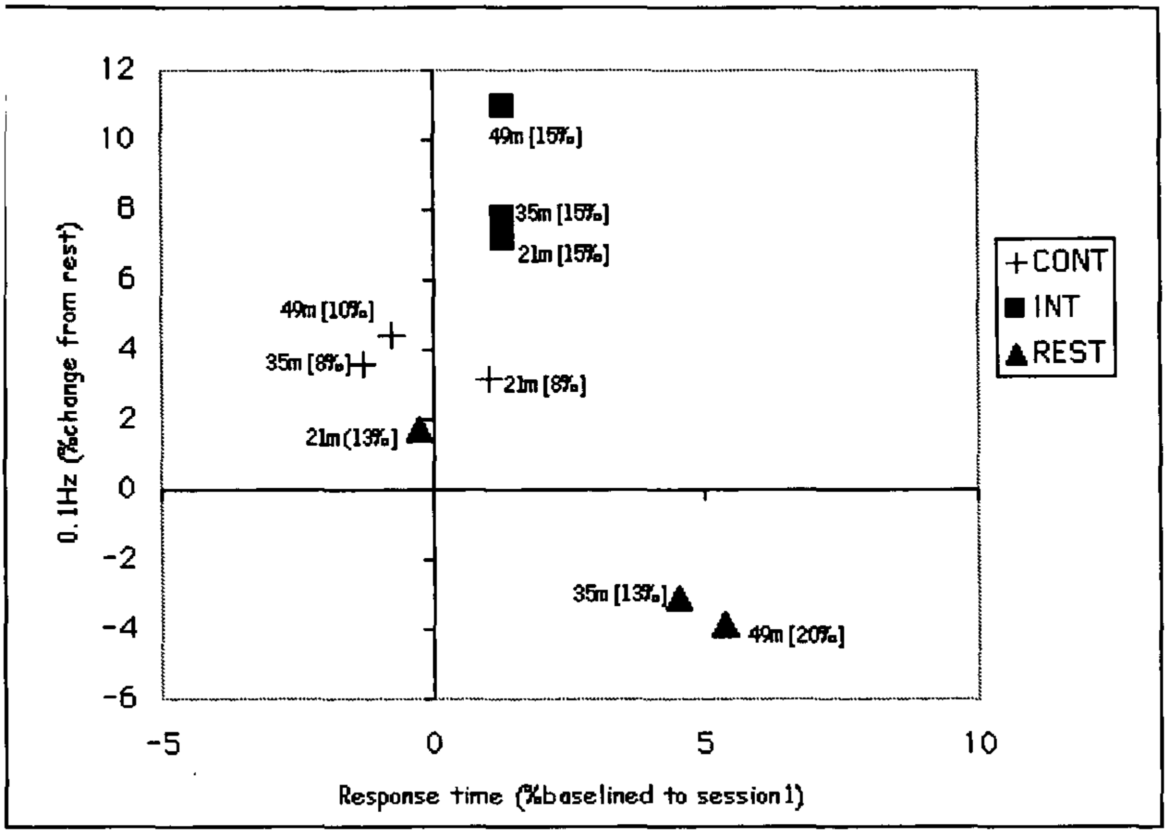


Figure 42. Performance efficiency index for reaction task performance (%baseline RT latency, %RT errors) and psychophysiological mental effort (%baseline power in 0.1Hz bandwidth) during PREDICTABLE tracking (N=24). NB: Positive x = increased mental effort from baseline, positive y = increased RT latency from baseline, the figures in parentheses represent % RT errors.

In Figure 42, it is apparent that the conservation of effort by the REST group resulted in increased response latency and high levels of RT error. By contrast, the CONT group were able to sustain levels of RT latency and accuracy, which were close to baseline values with only a modest level of effort investment. The high level of effort investment that characterised the INT group resulted in sustained response latency (relative to baseline), but a reduced level of response accuracy.

Response task data and psychophysiology from the unpredictable tracking condition were subjected to an identical analysis. The results are shown in Figure 43.

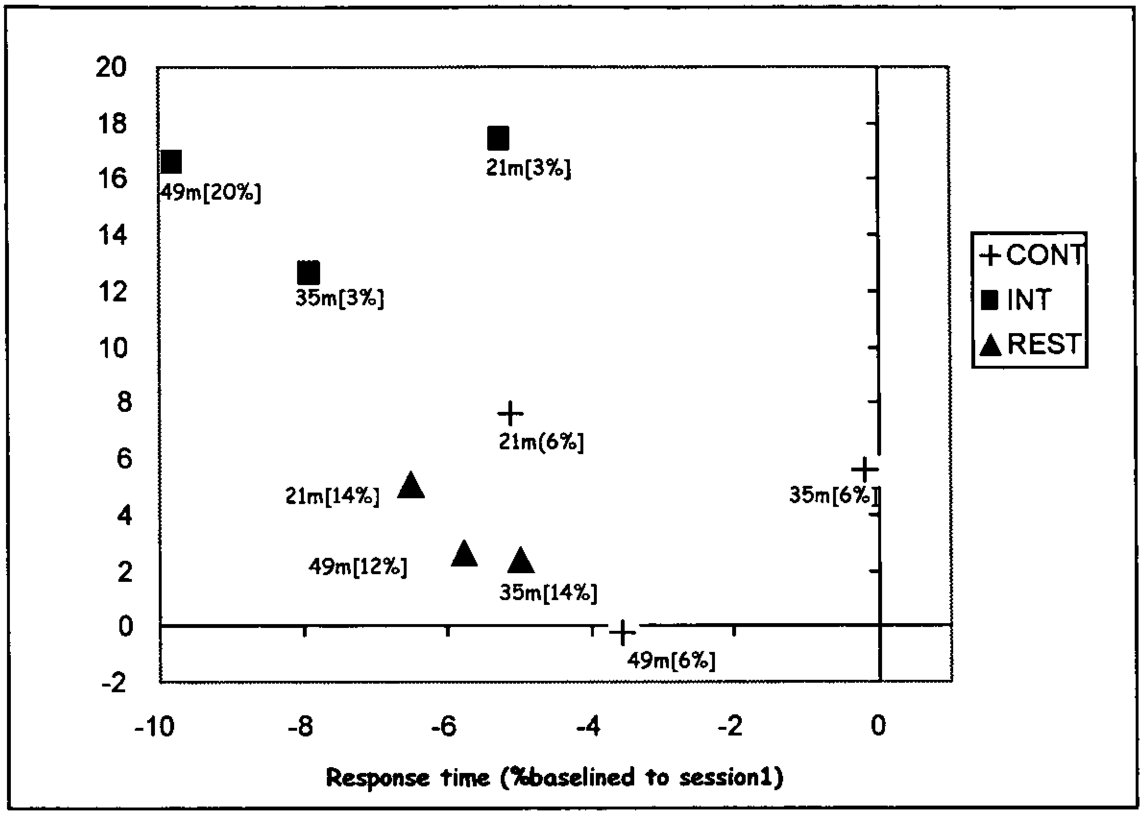


Figure 43. Performance efficiency index for reaction task performance (%baseline RT latency, %RT errors) and psychophysiological mental effort (%baseline power in 0.1Hz bandwidth) during UNPREDICTABLE tracking (N=24).

In this case, the pattern of results are similar to the predictable condition shown in Figure 42. The REST group conserve effort despite a consistent reduction of performance, i.e. increased latency and decreased accuracy. The INT group invested the highest level of mental effort overall, which resulted in low response latency and high accuracy until the final session, i.e. 20% decline of target accuracy following 49 minutes of performance. On the other hand, the CONT group invested less mental effort as TOT increased, but managed to sustain the highest level of RT accuracy and only a five percent increase of RT latency.

7.4.6 The regulation of mental effort

A regression analysis was performed to investigate how effort regulation was predicted by external and internal factors during each difficulty condition. It was predicted that external factors would be more influential when task difficulty was high, i.e. individual would focus on performance-related factors at the expense of internal factors. The obverse hypothesis would apply when task difficulty was low, i.e. internal factors emphasised at the expense of external ones.

These hypotheses were investigated via a series of regression analyses conducted on four 7min segments of each task difficulty condition. External factors were represented by the frequency of collisions whereas internal factors were indexed via a number of variables including; fatigue symptoms, energetic arousal, tense arousal and self-focus. These variables were entered as dependent variables with psychophysiological effort as an independent variable. Outliers were defined as subjects whose data was 2.5 standard deviations above or below the group mean.

The series of regression analyses performed on data from the Predictable task condition are summarised in Table 27 below.

Time elapsed	N	R^2	F-ratio	p
7min	21	0.64	5.64	<0.01
21min	22	0.70	7.63	<0.01
35min	21	0.42	2.16	n.s.
49min	21	0.57	3.89	<0.05

Table 27. Summary of regression analyses performed on data from Predictable task condition.

The semi-partial correlation for each predictor variable was calculated during the regression analyses. This coefficient indexes the amount of unique variance predicted by each predictor. The semi-partial correlation coefficients resulting from the regression analyses are illustrated below in Figure 44.

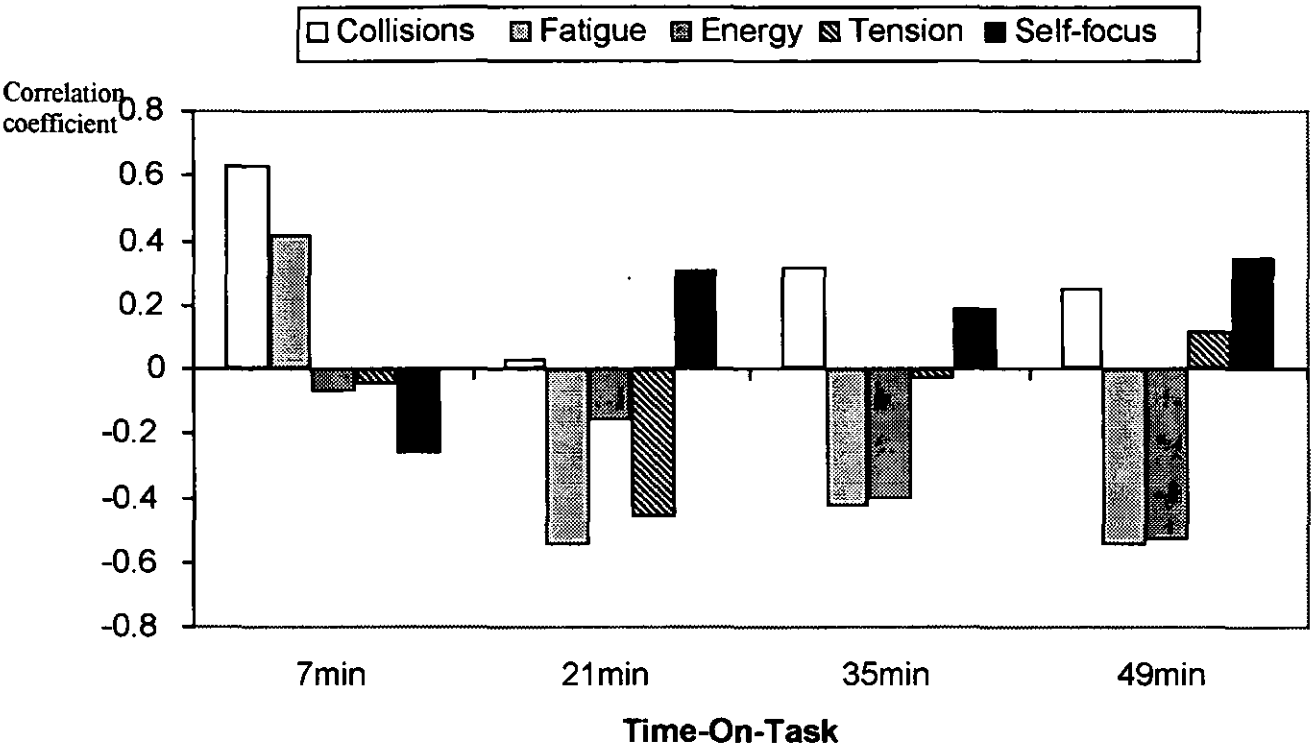


Figure 44. Semi-partial correlation coefficients from the regression analyses conducted on PREDICTABLE condition. NB: significant correlation > 0.40.

The data illustrated in Figure 44 revealed the variable relationship between mental effort and related factors. It is notable that an increased frequency of collisions was associated with mental effort investment. However, other variables such as fatigue symptoms and falling energetic arousal tended to lead to effort reduction or conservation. The relationship between effort and certain variables tended to change over time-on-task. During the initial period, rising fatigue increased mental effort but this positive association was reversed during subsequent periods of task activity. A similar reversal was noted for the level of self-focus, being negative during the initial period and positive thereafter.

An identical series of regression analyses was conducted on data from the Unpredictable task condition. The regression data is summarised below in Table 28 and the semi-partial correlations are illustrated in Figure 45. In this case, mental effort investment is driven by the number of collisions and reduced by the frequency of fatigue symptoms.

Time elapsed	N	R^2	F-ratio	p
7min	21	0.47	2.65	0.06
21min	21	0.71	7.27	<0.01
35min	22	0.61	4.95	<0.01
49min	21	0.66	5.73	<0.05

Table 28. Summary of regression analyses performed on data from Unpredictable task condition.

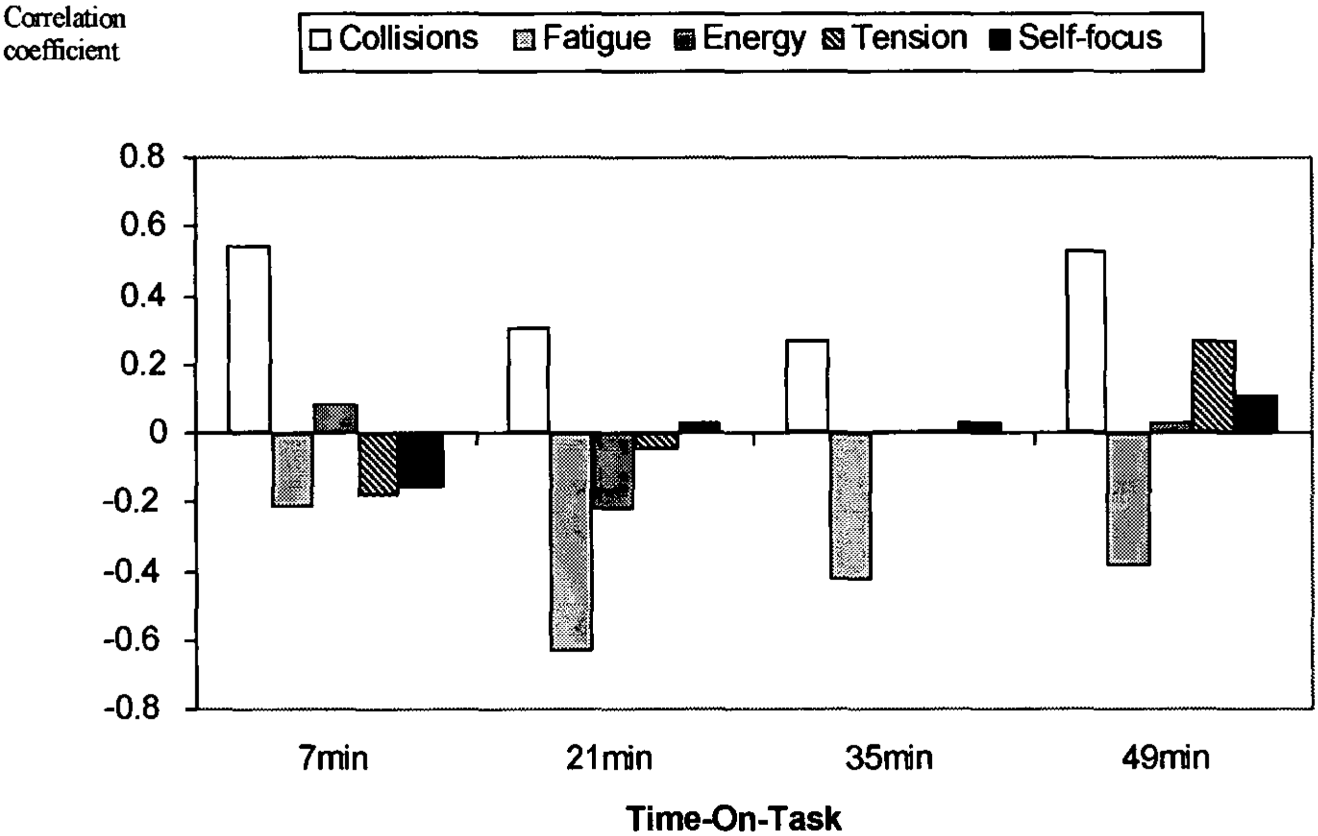


Figure 45. Semi-partial correlation coefficients from the regression analyses conducted on UNPREDICTABLE condition. NB: significant correlation > 0.40.

7.5 Discussion

The important experimental manipulations of tracking predictability and time-on-task (TOT) had the expected influence on performance, psychophysiology and subjective data. By escalating the unpredictability of the tracker target, perceptual-motor demands were increased, both with respect to tracking difficulty and the potential for interference between the tracking task and the target detection task. The net effect of target unpredictability was to reduce the accuracy of tracking performance, e.g. higher RMS error, more frequent collisions. However, increased timesharing demands between tracking and RT stimuli due to target unpredictability improved performance on the latter, i.e. target sensitivity increased and response times were shorter (Figure 37 and Figure 38 respectively). This effect may have resulted from the increase of psychophysiological effort during unpredictable tracking condition (Figure 39). According to this hypothesis, one influence of increased mental effort may have been increased timesharing efficiency between both task components, i.e. subjects 'break' from the tracking task to perform the RT stimuli for shorter periods as a means of protecting tracking performance during increased unpredictability.

The main experimental manipulation of increased tracking unpredictability exerted a range of influences on task strategy and subjective appraisal. In the first instance, subjects tended to reduce goal aspiration and to appraise declining performance standards during the unpredictable condition (Table 17 and Table 19). Despite a reduction of goal aspiration, subjects attempted to compensate for accelerated demands by increasing their subjective level of mental effort investment. In addition, subjective mental workload increased on all sub-scales of the NASA-TLX (Table 26). A series of costs were associated with mental effort investment during unpredictable tracking , e.g. increased symptoms of muscular fatigue (Table 20), reduced energy (Table 21), and increased distress (Table 22). It should be noted that the unpredictable tracking manipulation induced increased physical as well as perceptual/attentional task demands. This was evident from the higher level of mouse input, the greater number of muscular fatigue symptoms (Table 20) and the increased level of physical demand (Table 26) observed during unpredictable tracking.

The time-on-task (TOT) effect shared a number of influences with the tracking manipulation, e.g. degraded tracking performance, reduced energy and affective tone, increased mental demand and subjective workload. However, there was also evidence for a number of specific TOT effects. For example, it was apparent that psychophysiological mental effort declined with increasing TOT (Figure 39). This finding suggested either: (a) that mental effort is finite and investment could not be sustained across the whole 49 minutes of the task duration, or (b) that a learning effect occurred and subjects required lower consecutive levels of mental effort as TOT increased. This latter hypothesis is rejected on several grounds. In the first instance, by nature of the predictability manipulation, the opportunities for skill acquisition were very limited. The predictable tracking task was totally deterministic, whereas the unpredictable tracking manipulation produced totally random movement. In addition, the presence of a 12 -18 minute pre-test training session should have reduced the learning effect. Secondly, tracking performance did not improve over TOT, in fact, the opposite effect was observed with respect to RMS error, i.e. no evidence for improved performance with continued task exposure. Thirdly, subjects generally lowered goal aspirations and self-rated assessments of performance with increasing TOT rather than the opposite trend (Table 17 and Table 19) and finally, subjects perceived subjective mental workload and mental demand to rise with increased TOT (Table 26), i.e. task demands were heightened rather than being reduced by TOT. Therefore, it is proposed that the initial hypothesis is supported and mental effort appeared to be finite and declined over a sustained period, regardless of task difficulty.

It was apparent that subjects experienced accelerated costs due to both increased unpredictability and TOT. A number of costs were induced by both manipulations, e.g. reduced energy (Table 21) and affective tone (Table 22) coupled with increased muscular fatigue (Table 20) and increased mental demand (Table 26). Several other categories of cost were linked exclusively with effort investment induced by tracking unpredictability, e.g. increased tension and reduced confidence (Table 22), an increased frequency of task-relevant thoughts (Table 24) and increased subjective workload with respect to all sub-scales excepting mental demand (Table 26). It may be postulated that these costs function to provoke a strategy of mental effort investment. A third grouping of costs were associated with declining levels of effort due to TOT, these included: increased fatigue symptoms in all sub-categories excepting muscular fatigue(Table 20), reduced concentration (Table 21) and an increased frequency of task-irrelevant thoughts (Table 24). These latter costs may be associated with a strategy of effort conservation (Hockey, 1993; Hockey, 1997). Within Hockey's conceptualisation, effort conservation corresponds to an intentional strategy. The evidence for intentionality was mixed in the present case. On one hand, pre-task goal aspirations are reduced with increased TOT, however, there was no evidence for any influence of any systematic reduction of subjective effort mobilisation to parallel the decline of psychophysiological effort.

Those costs that were common to both main effects may be ambivalent with respect to mental effort policy. For instance, effort may be invested or conserved as a result of falling energy levels, increased mental demands or negative affect. In the former case, the individual may wish to increase mental effort in order to compensate for reduced energy or increased mental demands. Alternatively, the individual could choose to reduce mental effort in the face of diminished energy or heightened mental demands.

The evidence discussed so far clearly indicates that mental effort is increased or invested due to the influence of unpredictable tracking. It was also apparent that mental effort tended to decline or to be conserved with increasing TOT. This antagonistic relationship between these main effects is fundamental to the following discussion of the group effects.

The REST group were granted a seven-minute rest break between each period of task performance. The overall effect of rest break provision was to sustain those costs associated with effort investment whilst ameliorating the influence of those costs associated with effort conservation. For instance, subjects in the REST group reported an increased frequency of task-relevant thoughts (Table 24), higher concentration (Table 21) whilst experiencing a lower level of mental demand than others subjects (Table 26), regardless of tracker predictability. These findings suggested that frequent rest periods may preserve effort-preserving costs whilst inhibiting the accumulation of harmful costs.

For reasons of clarity and consistency, further effects due to the Group variable are discussed with reference to each tracking condition as this interaction was crucial to the schedule experienced by each individual subject Group.

During the predictable tracking condition, the CONT group experienced a sustained period of low-demand task activity, the INT group performed the same task between periods of unpredictable tracking, and the REST group performed the predictable tracking task between rest periods. Therefore, it would be anticipated that TOT effects may be most pronounced for the INT group. No evidence for this trend was apparent from the analyses of performance data. With respect to the response time task, the CONT group achieved a higher level of perceptual sensitivity relative to the INT group (Figure 37), but this advantage was dispersed by increasing TOT. There was also a surprising effect that a significantly longer response latency was associated with the REST group (Figure 37) following 35 minutes of the experimental session. This may have been indicative of rising complacency due to the low demand associated with the predictable tracking task.

The analysis of task efficiency shown in Figure 40 illustrated that both CONT and INT subjects produced similar levels of tracking performance. However, the INT subjects were forced to invest a higher level of mental effort in order to achieve this equivalence. The same pattern of data was observed for the RT performance, except that RT accuracy is slightly lower for the INT group (Figure

The evidence discussed so far clearly indicates that mental effort is increased or invested due to the influence of unpredictable tracking. It was also apparent that mental effort tended to decline or to be conserved with increasing TOT. This antagonistic relationship between these main effects is fundamental to the following discussion of the group effects.

The REST group were granted a seven-minute rest break between each period of task performance. The overall effect of rest break provision was to sustain those costs associated with effort investment whilst ameliorating the influence of those costs associated with effort conservation. For instance, subjects in the REST group reported an increased frequency of task-relevant thoughts (Table 24), higher concentration (Table 21) whilst experiencing a lower level of mental demand than others subjects (Table 26), regardless of tracker predictability. These findings suggested that frequent rest periods may preserve effort-preserving costs whilst inhibiting the accumulation of harmful costs.

For reasons of clarity and consistency, further effects due to the Group variable are discussed with reference to each tracking condition as this interaction was crucial to the schedule experienced by each individual subject Group.

During the predictable tracking condition, the CONT group experienced a sustained period of low-demand task activity, the INT group performed the same task between periods of unpredictable tracking, and the REST group performed the predictable tracking task between rest periods. Therefore, it would be anticipated that TOT effects may be most pronounced for the INT group. No evidence for this trend was apparent from the analyses of performance data. With respect to the response time task, the CONT group achieved a higher level of perceptual sensitivity relative to the INT group (Figure 37), but this advantage was dispersed by increasing TOT. There was also a surprising effect that a significantly longer response latency was associated with the REST group (Figure 37) following 35 minutes of the experimental session. This may have been indicative of rising complacency due to the low demand associated with the predictable tracking task.

The analysis of task efficiency shown in Figure 40 illustrated that both CONT and INT subjects produced similar levels of tracking performance. However, the INT subjects were forced to invest a higher level of mental effort in order to achieve this equivalence. The same pattern of data was observed for the RT performance, except that RT accuracy is slightly lower for the INT group (Figure

42). Further evidence for higher levels of mental effort investment was apparent from the pattern of costs shown by the INT group. In addition, the INT group reported increased extrinsic motivation (i.e. increased concerns regarding success and failure, see Table 24) and visual fatigue (Table 20) relative to the other two groups. It is hypothesised that the increase of mental effort and associated costs resulted from the accelerated influence of TOT due to the interpolation of the unpredictable task. Therefore, the INT subjects were forced to increased mental effort in order to compensate for 'carry-over' effects from the unpredictable tracking sequence, i.e. a fatigue after-effect, (Hockey, 1993; Hockey, 1997).

The CONT subjects experienced a consistent and highly monotonous level of task demand during the predictable tracking condition. On the sole basis of a mental effort model, it would be predicted that the energetical demands of this schedule were lower than those experienced by the INT subjects. However, it may be overly simplistic to characterise sustained, low-demand activity as relatively effortless. For example, (Warm & Dember, 1998) has argued convincingly that the costs associated with continuous signal/non-signal discrimination are substantive, due to time uncertainty and low levels of situational control. Several strands of evidence to support this position were apparent with respect to the CONT group. These subjects were distracted by a higher frequency of task-irrelevant thoughts (Table 24) and lowest level of confidence in their own performance (Table 21). Therefore, it may be reasonably assumed that a reduction of energetical demands with respect to TOT (relative to the INT group) did not reduce the level of task difficulty encountered by the CONT subjects.

It is hypothesised that the monotonous character of the CONT schedule during the predictable tracking condition may have inhibited mental effort investment for two reasons. On one hand, overt errors such as collisions were relatively rare events, so subjects were generally deprived of an objective source of performance feedback during the unpredictable condition. Therefore, CONT subjects were aware of accelerated workload, but this awareness failed to promote increased task engagement, or to challenge individuals with respect to performance capacity (i.e. Distress as defined by (Matthews, et al., 1997)), or to provoke self-evaluation of personal qualities or goals (i.e. Worry as defined by (Matthews, et al., 1997)).

In the case of the REST group, mental effort was also conserved during predictable tracking and actually fell below resting values (Figure 40 and Figure 42). This conservation strategy had mixed results on performance effectiveness, being associated with the lowest RMS error (Figure 40) and the highest RT latency/%error during later task periods (Figure 42). It could be hypothesised that the monotonous nature of the predictable tracking task had a soporific influence on subjects generally , which have been most apparent for those subjects who encountered the lowest level of task-induced stress (Dinges & Kribbs, 1991; Kjellberg, 1977a).

The unpredictable tracking condition represented high perceptual-motor demands. It was hypothesised that the influence of both condition and TOT would be maximised for the CONT group, who were required to sustain performance across seven consecutive task sessions. This schedule was less strenuous for the INT subjects who performed a predictable tracking session between each unpredictable sequence. Therefore, it was expected that the costs associated with TOT would accumulate at a lower rate for the INT subjects relative to the CONT group. Finally, the REST group had the easiest schedule where breaks were interpolated between episodes of unpredictable tracking.

It was apparent that overt performance showed the highest decline for the CONT group during the unpredictable tracking condition (Figure 36). In addition, response latencies produced by the CONT group significantly increased following 21 minutes of performance relative to the INT and REST subjects (Figure 38). However, an unexpected effect occurred with respect to perceptual sensitivity, where INT subjects produced the highest level of performance over 35 minutes, before slumping to the lowest sensitivity during the final seven-minute task period (Figure 37). This finding suggested that the INT subjects may have over-extended their capacity with respect to perceptual sensitivity following 43 minutes of performance and suffered a reversal during the final session.

The analysis of tracking efficiency provided a clear illustration of several effort strategies in action (Figure 41). The INT group showed the highest level of mental effort investment in combination with a relatively high level of tracking performance. This was an example of an effective effort investment strategy where the mobilisation of mental effort has a beneficial influence on performance. The REST group did not appear to mobilise mental effort to a significant extent (i.e. around 5% above rest), but were capable of sustaining tracking accuracy at the same level as the INT group (i.e. who were investing effort at a level between 14 and 18% above resting values). A combination of high performance and low mental effort defines an efficient effort conservation strategy, i.e. where good performance is achieved at minimal effort investment (Eysenck & Calvo, 1992; Schonpflug, 1986b). Therefore, it may be appropriate to label the strategy associated with the INT subjects as effective, but relatively inefficient compared to the REST group. The data from CONT subjects were also representative of a conservation strategy, but in this case, high performance was not apparent. In addition, the CONT group appear to progressively reduce mental effort across TOT at the expense of tracking performance. This conservation strategy may represent a rational reduction of performance standards as a means of reducing compensatory costs (Hockey, 1993; Hockey, 1997) – however, this strategy may have been implicit as there was no evidence of a conservation strategy from the CONT pre-tests of goal aspirations or the subjective effort scales.

The analysis of RT task efficiency indicated that CONT subjects sustained high accuracy at the expense of longer latency throughout the unpredictable tracking task (Figure 43). This finding suggested that CONT subjects may have compensated for poorer tracking by sustaining high RT accuracy. By contrast, the INT group decreased response latency, which had a detrimental influence on accuracy during the final task session.

It was apparent that an increased rate of effort expenditure was accompanied by a number of specific costs, e.g. increased subjective workload, stress, falling confidence and reduced cognitive interference. On the other hand, a decrease of effort expenditure (i.e. effort conservation strategy) was associated with increased symptoms of fatigue, reduced concentration and heightened cognitive interference. This distinction represents an indication that the presence of certain categories of cost may provoke increased expenditure, whereas others may have the opposite effect. A characterisation of costs or, rather, combinations of costs, working as 'activators' and 'inhibitors' received some support in the experiment. For instance, the combination of reduced Engagement with increased Distress, Worry and workload may have spurred the INT subjects to sustain effort investment during the predictable task. Whereas a combination of increased workload, decreased Engagement/Distress/Worry failed to do likewise for the CONT subjects. The same effect was apparent for the REST group who consistently experienced higher levels of 'activators' in combination with minimal levels of 'inhibitors.'

The influence of costs on effort policy would appear to be a question of degree rather than kind. For instance, the INT group faced with almost identical levels of costs as the CONT group during the unpredictable task were capable of sustaining mental effort investment. The key differences between both groups were arguably a higher rate of replenishment for the INT group and the lower level of tracking task error (but no evidence that INT subjects subjectively assessed performance to be superior to the CONT group). This and the previous example illustrates the difficulties of describing effort strategies, which are multifaceted in nature and multidimensional in character. It is apparent that effort policy is primarily influenced by performance efficiency and its associated costs. However, the lines of influence between efficiency and costs are cyclic in nature, i.e. efficiency is determined by costs and vice versa, and therefore causality is highly circular, which creates a host of problems for prediction.

It is also important to acknowledge that all subjects were fully aware of task demands and temporal schedule. The knowledge of task demand and remaining work duration may also influence the perception of task performance and associated costs.

Finally, it should be acknowledged that effort strategies may be highly idioscyncratic entities, subject to influences from trait as well as state variables ((Matthews, et al., 1997; Matthews, Schwean, Campbell, Saklofske, & Mohamed, in press). Therefore, the potential role of individual differences should be emphasised and subject numbers were minimal in the study to accommodate the experimental design. This factor is particularly problematic when one considers that one group sustained mental effort across TOT regardless of tracking predictability. This finding could have been predicted from the finite resource model of mental effort. Alternatively, the experimenter could have been unfortunate enough to have unwittingly included a small number of highly-motivated individuals within the INT group. For the most part, the results of the experiment are fairly consistent, but one would feel more confident in the data if the study were to be replicated with a more substantive number of subjects (i.e. N = 24 per group).

7.6 Conclusions

The aim of the experiment was to investigate how different schedules of task demand influenced effort strategies in conjunction with time-on-task. The manipulation of tracking unpredictability was intended to provoke mental effort investment, whereas increasing TOT was predicted to curtail investment or induce the inverse strategy of effort conservation. In broad terms, both manipulations were successful with respect to these hypotheses.

The beneficial influence of a schedule which included rest breaks was apparent, regardless of the level of task demand imposed on the subjects. The advantages conferred on the REST subjects included high tracking performance in combination with reduced mental effort (i.e. improved efficiency) and an increased willingness to set goals at a higher standard. The presence of interpolated rest breaks meant that accumulated harmful costs due to performance had sufficient time to disperse before the next period of tracking performance. Therefore, the REST subjects were able to devote mental effort solely o task demands, which resulted in superior levels of task efficiency (i.e. during predictable tracking, this group operated at levels of mental effort investment which fell below resting values).

A rate of expenditure hypothesis was supported in the predictable tracking condition, i.e. the highest rate of expenditure was associated with highest level of performance inefficiency. During the unpredictable condition, the highest level of inefficiency was associated with the lowest level of effort expenditure. The former relationship is indicative of a strategy of effort investment, whereas the latter corresponds to an effort conservation strategy.

It was difficult to predict effort policy based on those cost associated with expenditure, e.g. workload, fatigue, Engagement, Distress and Worry. These costs were capable of functioning as activators or inhibitors with respect to effort expenditure. When task demands were low, effort investment was associated with accelerated costs and reduced performance efficiency. When demands were high, both the highest and lowest level of effort expenditure were associated with a similar pattern of costs. It may be concluded that performance efficiency must be considered in conjunction with costs in order to characterise effort policy.

The results of the study provided a clear indication of the significance of mental effort policy, both in terms of the development of cognitive-energetical theory and as a variable for consideration within applied research on work/rest schedules within industrial settings.

8 DISCUSSION

This discussion section will deal with specific areas of the thesis in a modular fashion. The initial Section 8.1 will provide an assessment of major findings from the empirical work within the context of the effort regulation model described in Chapter 2. The following Section 8.2 will attempt to refine and to extend the model of effort regulation on this basis of experimental findings. The third section (Section 8.4) describes possibilities for future research and Section 8.3 will provide a critical assessment of the thesis.

8.1 Experimental findings

The empirical work started from the hypothesis that mental effort represented a finite resource for the modulation of sustained, perceptual-motor performance. This hypothesis was the basis of both the theory of mental effort regulation (Chapter 2) and the foundation for the following three experimental studies. Empirical evidence to support this position originated from three sources of data from the final laboratory study (Chapter 7):

- Those subjects permitted interpolated rest breaks achieved higher levels of performance and performance efficiency, regardless of the level of task demand. In addition, rested subjects did not accumulate energetical costs with respect to time-on-task.
- When the task was relatively undemanding, the interpolation of high demand activity accelerated the rate of energetical costs and reduced performance efficiency (Figure 40).
- When subjects were required to sustain performance on highly demanding activity, this group performed poorly (Figure 36), experienced high costs and exhibited reduced performance efficiency (Figure 41) relative to groups who were not subjected to the same level of sustained, high task demand.

These findings support the hypothesis that mental effort represents a finite commodity. In addition, it was apparent that effort expenditure has a linear relationship with psychological costs (e.g. lowered task engagement, increased distress and worry). Evidence was also found for a third hypothesis that sustained effort expenditure was associated with declining performance efficiency as indexed by (Meijman, 1995), i.e. the requisite level of mental effort investment necessary to sustain performance tends to increase with time-on-task. The experimental findings described in Chapter 7 clearly indicated that all trends of increased effort expenditure and accumulated costs were ameliorated by the provision of rest breaks.

The hypothesis linking effort expenditure with declining performance efficiency is particularly striking as all four studies showed a significant time-on-task effect for psychophysiological mental effort, i.e. effort tended to decline with sustained performance. Three working hypotheses may be generated on the basis of this finding: (a) a finite resource effect, i.e. the gradual exhaustion of a finite effort reserve, (b) a learning effect, i.e. effort level declines as subjects familiarise themselves with the task, or (c) a novelty effect, i.e. subjects initially invest high levels of effort because they are new to the task. The learning effect hypothesis is rejected because no other evidence for skill acquisition was apparent during the studies. In all four studies, the general influence of increased time-on-task was to reduce performance quality, reduce the assessment of performance quality and increase subjective workload. The opposite results would be anticipated if subjects acquired increased skill with task exposure. A novelty effect is more plausible as often effort investment during the initial task period is substantially higher than consequent periods, e.g. Figure 39. On the other hand, subjects received a substantial period of task familiarisation prior to data capture in at least two of studies, e.g. pre-task familiarisation was approximately 15 – 25 minutes for the study reported in Chapter 7. It is argued that even if a novelty effect were present, its influence would be relatively short-lived and could not explain the reduction of effort and performance that occurred during the latter periods of task performance.

It was anticipated that both increased task demand and accumulating time-on-task represented a dual 'drain' on finite effort reserves. The systematic manipulation of both variables in Chapter 7 yielded a pattern of results supporting the economic conceptualisation of finite effort being expended at different rates per unit time due to the main effects of task difficulty and time-on-task.

These finite boundaries on effort expenditure were explored in other studies described in the thesis. The study of individual differences (Chapter 4) revealed that less-proficient individuals tended to invest mental effort at a higher level than proficient counterparts. On this basis, it would be anticipated that the less-proficient would exhaust effort reserves at a higher rate and therefore, rapidly reduce their capacity for performance (Schonpflug, 1983) (Figure 22). In other words, these individuals with a higher rate of effort expenditure are more likely to encounter capacity limitations after a shorter period of activity and/or at lower levels of task demand.

The issue of a performance capability was reprised in the first study of driver impairment (Chapter 5) which yielded results that are more complex. This experiment manipulated the magnitude of stressor by including both partial- and full-sleep deprivation groups within the design. It was anticipated that sleep deprivation would lower the level of finite effort reserves and therefore, diminish the performance capability of both groups of sleep-deprived subjects. This hypothesis was supported by the behaviour of the full sleep deprivation group. These individuals performed poorly, exhibited performance inefficiency and reduced effort investment with increased time-on-task. It was assumed that capacity was reduced because effort reserves were decimated by sleep deprivation before performance.

The same logic could not be applied to the behaviour of those subjects who were partially deprived of sleep. It was assumed that partial-sleep deprivation would also reduce the quantity of finite effort reserves, but perhaps not to same extent as a full night without sleep. However, these subjects were capable of the highest level of sustained effort investment during the 120min task session (Figure 27). This finding cannot be explained by recourse to a simplistic fuel analogy or a one-to-one link between finite effort reserves and performance capacity. It is argued that this complex finding resulted from the transactional basis of the effort regulation mechanism (Section 2.2.6)and the appraisal of costs.

It was postulated that the subjective assessment of cognitive-energetic efficiency via external and internal feedback (Sections 2.2.4 and 2.2.5) acted as a conduit between effort reserves and performance capacity.

The evidence for external and internal feedback was provided by a range of subjective measures used throughout the experimental studies. It is hypothesised that several subjective scales were related to external feedback, e.g. workload scales such as mental/physical/temporal demands and performance. On the other hand, subjective scales associated with fatigue and stress (as defined as Engagement/Distress/Worry by Matthews, et al. (1997) were analogous to the process of internal feedback. Based on the final experiment (Chapter 7), it was suggested that specific patterns of cognitive-energetical appraisal were associated with an increased level of effort expenditure, e.g. increased workload/task-relevant thoughts/Distress, whereas an increase of fatigue symptoms/task-irrelevant thoughts and reduced concentration were associated with declining effort.

It is hypothesised that the magnitude of external/investment "costs" relative to the extent of those internal/conservation "costs" represents an operationalisation of efficiency assessment as described in

Table 4.

However, these process of subjective self-appraisal are associated with a degree of indeterminacy. This was illustrated during the first experiment (Chapter 4) which provided some evidence of how this process of cognitive-energetic assessment could go awry. In this case, low-proficiency individuals were insensitive to both internal (e.g. energy levels) and external (e.g. performance) feedback cues. This effect was particularly marked in the presence of sleep deprivation, lending some support to the hypothesis that impaired psychological functioning extends to those processes of self-monitoring as well as task performance (Brown, 1994; McDonald, 1987). Based on an erroneous appraisal of efficiency, low proficiency subjects did not attempt to mobilise subjective effort as a compensatory response to sleep deprivation (Figure 18). This pattern of data was replicated in the driving impairment study (Chapter 5), when subjects were provided with alcohol. Two psychological manifestations of alcohol are increased self-confidence and an accentuated sense of wellbeing (Walls & Brownlie, 1985). In other words, the intoxicated subject may experience a bias with respect to both external feedback (i.e. believing performance to be superior than it actually is) and internal feedback (i.e. being insulated against task-related stress and other sources of discomfort). Support for this hypothesis were manifested for the intoxicated group via the pattern of degraded performance (Figure 26) coupled to a decline of workload-related frustration (Table 7) and the absence of any perception of significant performance degradation.

Awareness and accurate appraisal of internal and external feedback is the foundation of mental effort regulation. This assertion is underlined by evidence that internal and external symptoms or cues may function in two distinctive modes. For example, feedback may function as a predictive indicator. This capability was demonstrated in the second study (Chapter 4), where correlational data suggested that proficient individuals employed internal and external cues to mobilise effort investment. These data illustrated a positive association between effort investment and decreased electrocortical arousal in one condition, indicative of a compensatory strategy, i.e. to invest effort in order to counteract drowsiness (Figure 23a). When operational conditions were more strenuous, the same individuals shifted the emphasis to external feedback as a means of protecting task performance. In this case, effort was invested in response to increased error frequency (Figure 24a).

The second mode of operation for feedback cues is an active impact on effort expenditure. In this case, internal cues distract the individual from primary task performance (and may reduce performance) unless additional effort is invested as a compensatory strategy. The four studies provided numerous examples of this phenomenon in action. For example, the interpolation of high-demand activity increased the level of Distress and Worry for subjects during subsequent periods of low-demand activity (Chapter 7). The influence of sleep deprivation extended beyond the obvious (e.g. high subjective sleepiness and low energy) to increase Worry (task-relevant thoughts), frustration, temporal demands and physical demands (Table 7). In the final study (Chapter 6), the rate of gain associated with subjective symptoms of fatigue was correlated with the rate of effort expenditure, i.e. higher gain rate of subjective symptoms = decreased effort expenditure per unit time. In addition, both factors were implicated in the subjects' decision to withdraw from the simulated journey (Table 12 and Table 13).

The levels of internal costs exert an influence on the assessment of cognitive-energetic efficiency in conjunction with external feedback, which concerns the quality of task performance. The study on performance feedback (Chapter 6) brought external feedback to prominence via the presentation of overt performance cues based on objective data. This manipulation prevented a deterioration of performance due to time-on-task without increasing the level of mental effort investment (Figure 35). This finding suggested that the deterioration of perceptual-motor performance due to time-on-task, which was observed in all four studies, may have been due to: (a) an inaccurate appraisal of performance quality that was counteracted via the provision of objective feedback, and/or (b) a reduced fidelity of performance checking, which was counterbalanced by the overt presentation of external feedback cues. In addition, the absence of any quantitative increase of mental effort in association with sustained performance effectiveness may indicate the significance of a qualitative aspect of procedural effort investment, i.e. the timing of effort investment as opposed to the quantity of effort investment.

There was some evidence from the performance feedback study that the presence of feedback accelerated an awareness of effort expenditure and subjective fatigue, with respect to the decision to discontinue the task (Table 12 and Table 13). This link provides some support for the hypothesis that an awareness of internal costs may be accentuated by overt cues, which are associated with external performance feedback (Figure 11).

The studies of simulated driving endowed the subjects with a increased degree of personal control, as they were able to manipulate the speed of the vehicle and therefore, exercise a degree of discretion with respect to task pacing. This aspect was apparent during the study of driver impairment (Chapter 5), where evidence for several latent decrements associated with performance were apparent (Hockey, 1993; Hockey, 1997). These decrements seem to represent stratagems to improve performance efficiency without reducing effectiveness. In this case, it was found that subjects who had been partially deprived of sleep permitted performance to decline to a sub-critical level, before affecting any corrective action. This strategy permitted the same subjects to reduce their level of steering input as an indirect result (Figure 26). These strategies did not appear to alleviate the level of mental effort required by the PartSD subjects to sustain performance (Figure 29) – which was not surprising, given that these subjects had to compensate for high internal costs provoked by a night of reduced sleep. On the contrary, the presence of these latent decrements hint of an increased awareness of external feedback and greater fidelity of control over performance, i.e. the ability of PartSD subjects to respond proactively to near-lane crossings. This hypothesis is also based on the relative difference between the PartSD group and the FullSD subjects (who reduced steering input but only at the expense of lateral control) and the intoxicated subjects (who provided a 'normal' level of steering input coupled to poor lateral control).

The experimental analyses from all studies included an analysis of performance efficiency based on the relationship between performance and psychophysiological mental effort. This formula replicated the analysis performed by Meijman (1995) and represented an operationalisation of performance efficiency as conceived by Schonpflug (1985) and Eysenck & Calvo (1992). Performance efficiency referred to the ratio between performance effectiveness and the level of effort required to achieve that level of effectiveness. It should be distinguished from the appraisal of cognitive-energetic efficiency, which represents the subjective utility between external and internal feedback cues.

The analysis of performance efficiency permitted performance quality to be assessed in conjunction with the level of effort investment. This two-dimensional representation was used to distinguish effective from ineffective effort policies. For example, the group who received rest breaks during the first study (Chapter 7) exhibited efficient investment, i.e. maximum performance in combination with minimum increase of effort (Figure 41), as well as a highly efficient effort conservation strategy, i.e. maximum performance in combination with reduced mental effort (Figure 40). The study of individual differences contrasted the performance efficiency of high and low proficiency subjects, following a normal night of sleep and a night without sleep. The analysis illustrated in Figure 21 clearly indicated that high proficiency was associated with increased performance efficiency. This distinction between both groups was particularly evident, as sleep-deprived, proficient subjects exhibited higher efficiency than rested, low proficiency subjects. During the study of driving impairment (Chapter 5), the same analysis indicated that partial-sleep deprivation decreased efficiency but not at the expense of performance, i.e. the partially-sleep deprived subjects increased effort investment to sustain equivalent performance to the control subjects (Figure 29). By contrast, the same analysis revealed that full sleep deprivation degraded performance as those subjects reduced the level of effort investment. The primary manipulation during the final study of sustained performance (Chapter 6) was the presentation of performance feedback. The analysis of performance efficiency indicated that discrete feedback sustained a higher level of performance efficiency relative to the control condition (Figure 35). In this case, the provision of feedback promoted increased performance effectiveness without inflating the requisite level of mental effort investment.

It is argued that these data justified the inclusion of the performance efficiency index as a means of operationalising successful and unsuccessful effort strategies within a two-dimensional framework. A successful effort policy may be characterised as an episode of investment that improves performance or an instance of effort conservation, which reduces the intensity of internal costs. Unsuccessful effort investment does not improve performance and an ineffectual conservation strategy would fail to reduce the intensity of internal costs (Figure 9). Performance efficiency data may be interpreted alongside the analyses of gain rates associated with psychological costs to represent the success or failure of effort conservation strategies. It was anticipated that the rate of gain would decline if effort conservation was successful. In the case of the subjects who received rest breaks during the final laboratory study (Chapter 7), these individuals actually experienced a negative rate of gain for most sources of internal costs, i.e. costs actually showed a progressive reduction with time-on-task. Examples of all four categories of effort policy were apparent from the four studies, based on the analysis of performance efficiency and the rate of gain associated with psychological costs.

The success or failure of an effort policy may have profound implications for the experience of psychological cost. For example, in the study of driver impairment (Chapter 5), partial-sleep deprivation subjects experienced a similar level of cost to those who had been fully sleep deprived. However, the investment strategy enforced by the former group successfully counteracted the influence of sleepiness on performance. Whereas the latter group experienced a combination of performance loss and increased costs, which forced them to reduce their level of mental effort. These results indicated that those internal, psychological costs, characterised as discomfort, may be tolerated in combination with a successful task performance. This example illustrates the transactional relationship between effort, costs and performance.

A dissociation between subjective and psychophysiological indicators of mental effort was observed in two of the four studies. As stated earlier, the standard influence of time-on-task on the 0.1Hz component of heart rate variability was a systematic reduction of mental effort. However, the subjective effort scale (from the Raw version of the NASA-TLX) indicated that subjects perceived their level of mental effort to increase with rising time-on-task (Table 26 and Table 7). This dissociation may be indicative of a subjective bias (i.e. the need to feel as though we are working as hard as possible) or it may simply reflect the fact that subjective feelings of effort mobilisation (i.e. our intentional effort policy) may dissociate from actual effort policy. This topic will be revised in the following section.

These findings represent the core results from the empirical programme, the following section will attempt to reconcile experimental findings with the model of effort regulation described in Chapter 2.

8.2 Theoretical issues

The results of the four experimental studies produced some contradictory indications concerning the finite limitations of mental effort investment over sustained time-on-task. It was postulated in Section 1.2 that finite effort reserves could be characterised according to an effort economy. In addition, it has been suggested that the finite character of mental effort reserves may be physiological in nature, i.e. being related to the efficiency of cerebral metabolism, and in particular, to energy consumption at the synapse (pumping out ions to prepare the synapse for subsequent signals). It was reasoned that the failure of the latter process would reduce the fidelity of synaptic transmission (Crawford, 1961; Tsaneva & Markov, 1971), which would represent an upper limit on performance effectiveness. This tenuous line of theorising links effort to energy and energy to the fidelity of nervous transmission, which in turn, has been assumed to influence performance. Moreover, this framework posits the existence of an absolute, physiological limit on mental effort expenditure over time.

In broad terms, the experimental data supported a finite conception of mental effort. There was evidence that the rate of effort expenditure was influenced by the level of task demand and time-on-task in contrary directions. In addition, continued effort expenditure was associated with increased psychological costs and latent decrements with respect to performance. The empirical data indicated that performance efficiency was the primary influence on the rate of effort expenditure, i.e. expenditure being determined by the minimum requisite level of effort investment necessary to sustain adequate performance. Furthermore, it was apparent that psychological costs generally reduced the level of performance efficiency, i.e. increased costs = increased investment due to the requirement to compensate for task-irrelevant sources of effort consumption such as discomfort, distraction, reduced engagement etc. Therefore, the evolution of costs reduces the maximum level of performance efficiency, which raises the requisite level of effort expenditure whilst degrading the level of performance effectiveness. This scheme is represented schematically in Figure 46.

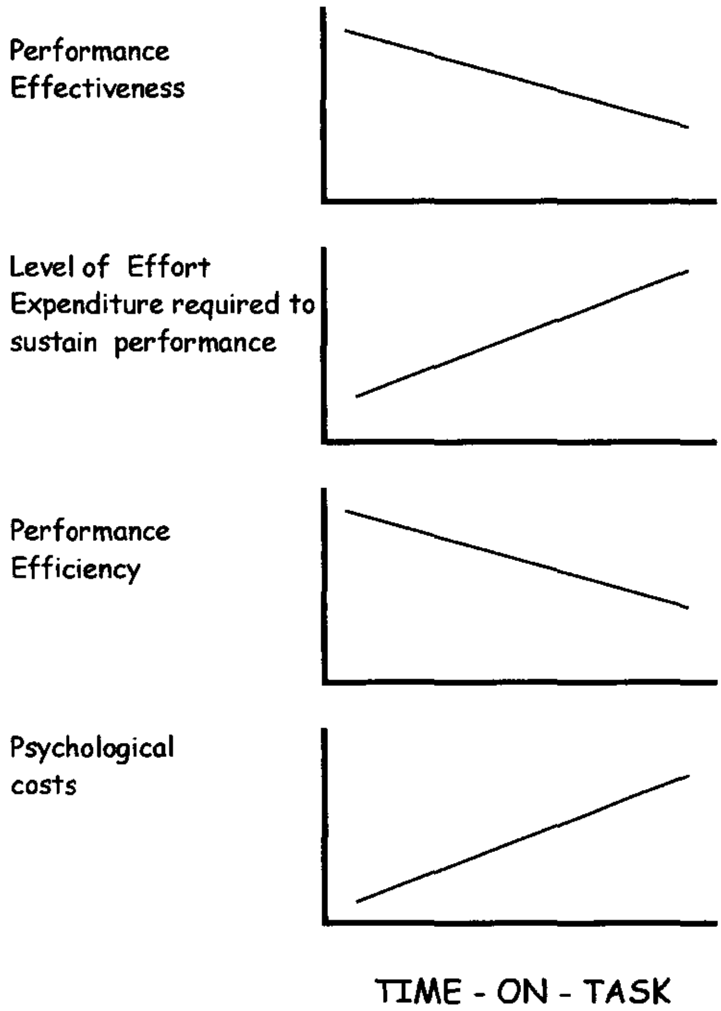


Figure 46. Schematic representation of the relationship between performance effectiveness, effort expenditure, performance efficiency and psychological/energetical costs.

It was assumed that sleep deprivation would reduce the level of finite effort reserves available, and

:herefore, limit the maximum level and duration of effort expenditure. However, the evidence from Chapters 4 and 5 was contradictory with respect to this claim and begs a number of questions concerning the assumption of absolute limits on effort expenditure based on physiological processes. This hypothesis is challenged by the alternative formulation of flexible limits on mental effort, based on psychological appraisal.

The notion of flexible limits on mental effort originates from the model proposed by Kahneman (1973). This author hypothesised that the level of effort available for task performance was determined by the level of psychological arousal, i.e. extremes of fatigue and stress inhibited the availability of mental effort. This model was contradicted by the results of the driver impairment study (Chapter 5). Both groups deprived of sleep invested mental effort at a higher level than the fully-rested Control group. Therefore, an alternative formulation is proposed where the amount of effort available for performance is determined by efficiency, valence and self-belief.

This hypothesis is based on the assumption that humans may only assess the level of effort expenditure/replenishment indirectly, via the cognitive-energetic assessment of behavioural efficiency. Those external sources of feedback related to performance effectiveness may indicate sub-optimal effort investment via increased errors or a rising perception of task demands. Internal feedback of psychological costs are analogous to the level of discomfort experienced by the individual, and as such, denote the need for rest and restitution. It is impossible for individuals to register cerebral metabolism, synaptic efficiency or any alternative manifestation of psychophysiological energy directly. Indirect and subjective processes of assessment may be the sole means of indexing remaining effort reserves.

It is hypothesised that flexible limitations on effort investment originate within this subjective framework of appraisal. For example, individuals may be capable of tolerating a high degree of discomfort in combination with good or acceptable levels of performance. However, it is anticipated that few are capable of sustaining effort investment in combination with task failure. This distinction was apparent in the study of driver impairment between the two sleep-deprived groups (Chapter 5). Those who had no sleep prior to test session experienced increased errors (lane crossings) with high discomfort (high sleepiness and workload). On the other hand, the PartSD group tolerated a similar level of discomfort because task performance was successfully protected from the influence of sleepiness. Therefore, the PartSD group was capable of sustaining a higher level of mental effort investment that the FullSD group. The flexibility of effort limitations may permit performance beyond what may be considered the 'normal' boundaries of endurance.

The proposition that the finite limits of effort investment are flexible and based on subjective appraisal should not be taken too literally. This hypothesis does not imply individuals may invest effort for an unlimited periods, based solely on subjective processes of appraisal. The perspective adopted by the model of regulation is identical with action theory, i.e. that appraisal involves a comparison between a subjective standard and an external source of feedback (Frese & Zapf, 1994). Therefore, the capacity of the individual to resist the objective reality of poor performance or the subjective presence of discomfort is bounded. For example, if a sleep-deprived operator continues to perform beyond a certain period of time, the symptoms of sleepiness will impose themselves on behaviour, resulting in performance lapses and microsleeps (Dinges & Kribbs, 1991). These limitations will degrade efficiency as a means of preventing task continuation. Under extreme circumstances, these limitations may simply induce sleep, suspending behaviour and the process of efficiency assessment alike.

It is proposed that flexible limitations (defined by efficiency, goals and self-belief) may account for the differences between high and low proficiency subjects within a sustained performance scenario (Chapter 4). Similarly, this process of subjective appraisal may have influenced the subjects' decision to discontinue the experimental task (Chapter 6).

The finite limitations on effort investment are proposed to result from the presence and magnitude of external and internal sources of feedback. It is postulated that external and internal feedback cues function as respective activators or inhibitors of mental effort investment. For example, the perception of increased task demands, increased error frequency, goal valence may activate higher levels of mental effort investment. On the other hand, rising symptoms of fatigue, Distress or Worry and reduced task Engagement may inhibit effort investment. The dual-appraisal of efficiency represents the crucial antagonism between the need to protect performance and to desire to protect the self from increased or continued discomfort.

The presence of activators or inhibitors influence the assessment of efficiency, however, both activators and inhibitors may also be associated with a degree of severity. Feedback from external sources may differ with respect to the severity associated with a particular cost, i.e. feedback indicating a slight error vs. catastrophic failure. A similar hierarchy of severity is apparent for those sources of internal feedback. For example, low levels of subjective fatigue and reduced task Engagement may be tolerated. However, if subjects experienced greater Distress (Matthews, et al., 1997), these symptoms (reduced affect and concentration, increased stress) indicates that the capability to perform is under threat. Furthermore, if the level of Worry increases – this may indicate that the individual is reassessing their self-knowledge with respect to personal qualities and capabilities (Matthews, et al., 1997). This latter category represents the most severe type of internal cost. This hypothesis was supported by the results of the final study (Chapter 7) and the hierarchical analysis of internal costs is based on the S-REF model (Wells & Matthews, 1994) (Section 2.2.5).

A second question is raised concerning the process of cognitive-energetical appraisal of efficiency. Specifically, is external feedback appraised then buffered for subsequent comparison with internal feedback? Alternatively, does efficiency appraisal involve a simultaneous appraisal of both sources of feedback? According to Duval & Wicklund (1972), appraisal may oscillate between the self and external sources of information. However, this possibility was challenged by Wells & Matthews (1994), who postulated co-occurrence, i.e. a resource view wherein the distinction between self-focus and task-focus is determined by the relative proportion of attentional resources directed towards appraisal of the self and the task. An intermediate approach was suggested by Ingram (1990), who described internal/external appraisal in terms of the relative frequency and duration of self- and task-directed appraisal, i.e. the proportion of simultaneous appraisals directed to the self and the task. The data from the second study (Chapter 4) indicated that proficient subjects tended to regulate effort as an exclusive response to either external or internal feedback sources. These data represent preliminary findings but suggest a serial, switching strategy for feedback appraisal along the lines suggested by Duval & Wicklund (1972) and Ingram (1990). This serial strategy is supported by the model of goal management described by Baars (1988). According to this framework, goals are managed in a serial, sequential fashion. Therefore, a switching strategy for the appraisal of efficiency may reflect the relative prioritisation of: (a) the goal to protect performance, and (b) the goal to protect the self (i.e. to reduce discomfort).

One implication of the serial strategy is the prioritisation of external and internal feedback. A weighting of one direction of feedback at the expense of the other could lead to an incomplete or distorted assessment of task efficiency. For example, an individual who is highly motivated to succeed may prioritise external, performance-related feedback at the expense of internal symptoms of discomfort. This is a potentially hazardous strategy, as continued neglect of internal feedback may result in stress-related, health problems. At the other extreme, an individual who is prone to ıypochondria may find it difficult to attend to external feedback, because he or she habitually ›rioritises feedback from internal, physiological sources.

The original model of mental effort regulation (Figure 11) postulated a bi-directional link between external and internal feedback. This link permitted internal discomfort to accentuate task demands and vice versa. The purpose of linkage between external and internal feedback may represent a corrective mechanism to ameliorate the kinds of distortion of efficiency appraisal due to attentional prioritisation. Therefore, the work-obsessive described in the previous paragraph may be forced to acknowledge internal feedback, as it impacts on task performance. Similarly, the hypochondriac is informed that an acceleration of internal costs may be triggered by poor task performance. Support for this hypothesis was evident from the study of performance feedback (Chapter 6), when the presence of feedback appeared to raise awareness of subjective fatigue, which subsequently persuaded some subjects to discontinue the simulated journey.

This discussion of hierarchical costs and feedback prioritisation is intended to illustrate the complexities associated with mental effort regulation. The key to effort regulation and the formulation of policy is awareness of cognitive-energetic efficiency at both the strategic and the procedural level. The experimental programme provided examples of unresponsive effort policy due to lack of awareness (Chapter 4), and insensitive effort policy due to the influence of biological stressors (Chapter 4 and 5). It is also hypothesised that a poor internal model of the task would lead to a distortion of external feedback, which may have a negative influence on effort policy (Section 2.2.4).

It was evident from the experimental work that psychophysiological and subjective indices of mental effort tended to dissociate with time-on-task. When effort was invested in response to task demands (Chapter 7), both indicators were in agreement. However, it was apparent that subjects perceived their level of effort to increase in order to compensate for time-on-task, whilst the opposite trend was observed in the psychophysiological variable. Several explanations are possible: (a) psychophysiological effort is influenced by a physiological confound, which decreases with time-on-task, (b) subjective assessment is insensitive to the gradual reduction of psychophysiological effort, or (c) subjective estimation reflects the intentions of the individual rather than the psychophysiological policy. The first hypothesis is rejected on the grounds that psychophysiological effort is associated with declining performance quality, which would be anticipated if the former acted as a modulator of performance. It is argued that both (b) and (c) provide explanations that are more satisfactory and may function in conjunction. For example, if the gradual reduction of effort goes undetected and the subjective estimates of the individual only represent an attempt to mobilise effort, i.e. an intention rather than an eventuality.

These findings call into question the link between effort investment and volitional control (Section 1.1). Those subjects who participated in these experiments desired to increase mental effort, but failed to do so according to psychophysiological data. This dissociation between subjective intentions and effort policy may indicate that the latter is formulated on either an active or a passive basis. For example, if we become absorbed in a task, we may invest substantive effort in the absence of any conscious intention (Kahneman, 1973). This scenario could be termed passive effort investment. On the other hand, as observed in the experimental data, increased time-on-task may reduce the level of effort investment due to internal costs without conscious realisation. This policy may be termed passive effort conservation. The active versions of both policies are linked to volitional control, i.e. investment to protect performance, conservation to ameliorate costs (Hockey, 1993; Hockey, 1997).

These theoretical topics provide an indication of the complexity associated with mental effort regulation and highlight certain limitations of the current document, which are described in the following section.

8.3 Limitations of the thesis

The diverse nature of the thesis topic necessitated a broad approach to literature review and theory generation. This expansive approach collated diverse areas of psychological research under a common conceptualisation, and as such, was associated with a number of short-comings. For instance, the model of effort regulation is based on the fundamental hypothesis that effort is capable of modulating the quality of performance.

The model of mental effort regulation rests on a second fundamental hypothesis that mental effort is finite. The empirical work provided several examples of indirect support for this hypothesis, involving psychophysiological indices such as the 0.1Hz component of sinus arrhymia. However, these indirect sources of data fail to resolve the finite reserve or fuel metaphor for mental effort. A more direct source of potential data is represented by recent developments in brain imaging techniques (Haier, et al., 1988), where cognitive functioning may be indexed directly to processes of cerebral metabolism in the brain. It was not possible to address this issue of linkage between cerebral metabolism, effort and performance in the current document.

The proposed model of mental effort was relatively simple, involving dual feedback loops to a 2-level hierarchy for effort regulation. However, this simple model produced a high number of possible patterns and contingencies between performance, effort and costs (Figure 46). These patterns of association and dissociation may be characterised as qualitative states (Hockey & Hamilton, 1983). However, it is difficult to make concise predictions concerning performance quality based on various contingencies. Furthermore, it is possible to expand the framework for mental effort regulation to include other important constructs such as self-efficacy, goal standards, mental models of task etc. It is argued that the inclusion of such constructs would make the business of prediction more opaque. This is a problem which besets all psychological models to a greater or lesser extent, i.e. to be sufficiently inclusive yet capable of concise prediction concerning behaviour. In the case of mental effort, this tendency may be amplified by the diversity of psychological concepts associated with regulation and control.

It may be argued that operationalisation of the model placed too much emphasis on subjective, self-report techniques. For example, data from subjective questionnaires represented both processes of feedback and appraisal from external and internal sources. However, the effort regulation model emphasises the argument that both exist as separate stages in a single process.

It was also apparent that the operationalisation of variables was highly selective. For example, psychophysiological mental effort was represented by the 0.1Hz component. However, there were other potential candidates such as pupil diameter (Beatty, 1982b) or the P300 component of an evoked cortical potential (Ullsperger, Metz, & Gille, 1988). The representation of psychophysiological effort via a single variable also disregards the possibility that mental effort may be sensitive to phasic or tonic variables or exist as a multidimensional construct.

The empirical research performed within the thesis was limited in a number of important respects. In the first instance, the process effort regulation was presumed to reflect a large amount of individual differences, due to expectancies, past experience etc. It was expected that the study of such an idiosyncratic phenomenon would necessitate higher subject numbers than those reported in Chapter 4. Only the study of driving impairment (Chapter 5) approached the requisite number of subjects required for that particular experimental design. This flaw was particularly apparent for the between-subjects designs used in Chapters 4 and 7. This deficit is underlined by a consideration of the number of different measures employed in each study. Mental effort regulation requires the inclusion of multidimensional measures, however, the net result of performing a large number of statistical tests over many measures with low N is to accelerate the risk of a Type I error.

It is acknowledged that the thesis struggled to operationalise the dynamics of the effort regulation model with respects to statistical analysis. For example, a regression analysis of subjective/objective aspects of external/internal feedback would have been more appropriate than the correlational analysis to investigate cues for effort investment in Section 4. In addition, it was difficult to test those metrics that captured salient aspects of effort model for statistical significance, such as performance efficiency and the rate of gain of costs.

Several limitations of the current work constitute potential avenues of investigation during future research which are described in the following section.

8.4 Future research

The empirical work described in previous chapters failed to provide systematic coverage of the hypotheses arising from the model of effort regulation (Figure 11). This inadequacy was due to the largely applied nature of the experimental work. It would have been ideal to have tested and established a model of effort regulation in the laboratory with simple tasks before proceeding to ecologically valid, complex tasks such as driving behaviour. Therefore, the future research proposed below is generally concerned with basic issues which passed unexplored in the current document. This strategy may be seen as a step backwards, however, it is argued that the model testing must develop from rigid control of experimental variables. Several lines of future research are described in the following sections.

- The model of mental effort regulation is justified by an argument that effort represents a finite resource. This claim raises at least two problems: (a) work remains to be done concerning the operationalisation of the effort concept, and (b) the resulting effort variables must be studied within the realm of sustained performance where exhaustion of a finite resource remains a possibility. The former problem would constitute a doctorate thesis in its own right and was not the focus of the current document. It is postulated that finite boundaries on effort investment will be physiological in nature. Therefore, it is important to wed existing experimental work on effort operationalisation using heart rate variability, pupilometry etc. with contemporary advances in brain imaging and the measurement of cerebral metabolism. It seems reasonable to propose task-related effort investment represented by psychophysiology should have a relationship with an index of cerebral metabolism such as glucose uptake or blood flow. This relationship could be studied by adopting a similar design to that used in experiment 4 (Chapter 7), as an interaction between task demand and time-on-task.

- A second methodological point concerns the operationalisation of effort as heart rate variability used throughout the experimental work. As stated earlier, psychophysiology was not the focus for the current thesis, however, a range of candidate variables such as P300 and pupil diameter was also available. The inclusion of alternative indices of mental effort may have improved the sensitivity of psychophysiological effort. In addition, it may be useful to consider a multiple-resource conception of mental effort composed of 'visual effort' (pupilometry) and 'central processing' effort (P300). This would permit performance efficiency to be characterised with reference to more than one psychophysiological index. For example, the quantification of performance efficiency for a visual RT task shown in Figure 42 and Figure 43 could use pupilometry rather than 0.1Hz
- The initial laboratory study focused on individual differences as categorised by a post-hoc division based on task performance. This analysis indicated that poor performance resulted from high effort investment coupled with insensitive regulation. This line of investigation warranted further experimentation, but limited subject numbers in the following three studies curtailed this possibility. This was unfortunate because the model of effort regulation may benefit from the use of individual differences as an experimental manipulation. For example, it may be argued that neurotic individuals have a heightened awareness of energetical costs, and therefore regulate effort in line with internal as opposed to external feedback. Similarly, individuals with high self efficacy are problem/task-oriented and may exhibit the opposite bias during effort regulation. This line of investigation may have implications for the study of occupational stress.
- The driving impairment study (Chapter 5) suggested that intoxicated subjects performed poorly because they were less sensitive to external feedback of error. This hypothesis could be tested by providing intoxicated subjects with performance feedback as used in Chapter 6, to investigate if this manipulation resulted in an improvement of performance effectiveness and an increased awareness of internal costs.

- The final study described in Chapter 7 manipulated task demand and temporal schedule in order to influence the relative rates of effort expenditure and replenishment. This methodology could be replicated by varying the levels of demand, temporal scheduling and time-on-task, in order to investigate the adaptability of effort policy. If a larger range of test conditions could be generated, it may be possible to qualify various stages of effort policy, from high efficiency to extreme inefficiency, with respect to performance, psychophysiology and subjective data.
- The subjects who participated in the study of demand and time-on-task (Chapter 7) received full instructions prior to performance, concerning the duration and level of demand they would experience during the experimental session. It would be interesting to investigate how effort policy is influenced by expectations by not providing information on the task schedule prior to performance. It would be anticipated that subjects may be less inclined to sustain effort investment in the face of uncertainty.

9 CONCLUSIONS

Evidence from literature review and empirical study was gathered to support the hypothesis that mental effort is a significant construct for sustained, perceptual-motor performance. It is concluded that mental effort has an important theoretical position, as it constitutes a suitable candidate to unite cognitive and energetical (i.e. biological) traditions within psychology. In addition, it was argued that mental effort was relevant for applied research related to work/rest duration and safety-critical performance.

It was postulated that mental effort may modulate perceptual-motor performance via one of two routes: indirectly via top-down strategic adjustments of goal-setting (i.e. goal aspiration) or directly by responding to bottom-up discrepancy detection at the procedural level of perceptual-motor performance. In the former case, the investment of mental effort is concerned with those processes of decision-making, which exert a top-down influence on actual behaviour, i.e. goal-setting. The latter case of procedural effort investment involves the direct modulation of cognitive processes within the perceptual-motor chain. The precise foundations of mental effort could not be defined on the basis of the thesis. It is speculated that mental effort is related to cerebral metabolism at the physiological level and controlled processing at the psychological level.

The regulation of mental effort is based on a cognitive-energetical assessment of behavioural efficiency, i.e. an dual appraisal of performance quality and energetical costs. This act of appraisal is accomplished with reference to external feedback from task performance and internal feedback from the self. It should be noted that efficiency appraisal is characterised as a subjective assessment and therefore, is prone to bias and distortion from a variety of sources.

The formulation of effort policy is based on the appraisal of cognitive-energetical efficiency and represents various modes of cognitive-energetical coping. If efficiency is sub-optimal due to poor task performance, effort may be invested via either a strategic or procedural route to improve performance. If energetical costs are high due to discomfort, the individual may actively reduce effort investment to reduce the intensity of energetical costs.

It was concluded that the period of maximum effort investment was constrained by flexible psychological limitations, operating within the boundaries of physiological limitations. Therefore, the amount of effort reserves which could be made available for performance were determined by efficiency and goal-setting operating within the context of self-belief and self-knowledge.

The empirical research supported a number of hypotheses related to mental effort regulation, these are listed as follows:

- A reduction of effort expenditure is often associated with declining performance effectiveness.
- The rate of effort expenditure determines the maximum duration that error-free performance may be sustained.
- The level of performance efficiency (ratio of mental effort to performance) determines the rate of effort expenditure
- As effort reserves are expended, a number of psychological costs (e.g. task stress) and latent decrements (e.g. changes in performance quality) may be apparent
- The negative consequences of effort expenditure (i.e. degraded performance efficiency, accelerated costs or decrements) may be counteracted by withdrawal from task activity and restitution
- The presence of psychological costs have the potential to degrade performance quality
- If costs are present, additional effort investment is required to protect performance, which degrades performance efficiency and increases effort expenditure.
- The effort economy (physiological limitations on effort expenditure) are appraised via self-assessment of external performance and internal state

The awareness of performance feedback and internal feedback was essential for efficient effort regulation. Experimental findings indicated that sleep deprivation and alcohol reduced awareness of internal/external feedback, which affected effort policy and the quality of performance. In addition, the provision of performance feedback was found to sustain performance quality without increasing the level of effort expenditure. This finding illustrated the importance of external feedback as a cue for procedural effort investment.

The thesis model was constructed based on three related hypotheses (Section 1.1). The postulation of strategic and procedural effort within a regulatory framework represented an attempt to reconcile these hypotheses within a unified framework. Based on this research, it is concluded that the relationship between effort and performance effectiveness was ambiguous (hypothesis 1). The investment of mental effort may only improve performance if it is expended at a requisite level and directed to an appropriate site of action. Mental effort may be invested in response to other variables, such as psychological costs (hypothesis 2). It is concluded that costs represent an important source of internal feedback, capable of hindering performance at a high levels and increasing the requisite level of mental effort investment). It was hypothesised that an individual may decide whether or not to continue task performance on the basis of external feedback in combination with the frequency and severity of psychological costs. This chain of activity represents the final link between performance, costs and volitional activity (hypothesis 3).

10 REFERENCES

Aasman, J., Mulder, G., & Mulder, L. J. M. (1987). Operator effort and the measurement of heart rate variability. *Human Factors*, **29**(2), 161-170.

Allport, D. A. (1980). Attention and performance. In G. Claxton (Ed.), *Cognitive Psychology - New Directions* (pp. 122-153). London: Routledge and Kegan Paul.

Allport, D. A. (1992). Attention and control: Have we been asking the wrong questions? A critical review of 25 years. In D. E. Meyer & S. Kornblum (Eds.), *Attention and Performance XVI* (pp. 183-218). Cambridge, MA: MIT Press.

Allwood, C. M. (1984). Error detection processes in statistical problem solving. *Cognitive Science*, **8**, 413-437.

Anderson, G. H. (1981). Diet, neurotransmitters and brain function. *British Medical Bulletin*, **37**(1), 95-100.

Baars, B. J. (1988). *A Cognitive Theory Of Consciousness*. New York: Cambridge University Press.

Baars, B. J. (1992). A new ideomotor theory of voluntary control. In B. J. Baars (Ed.), *Experimental Slips and Human Error: Exploring the Architecture of Volition* (pp. 93-120). New York: Plenum.

Baars, B. J. (1993). Why volition is a foundation issue for psychology. *Consciousness and Cognition*, **2**, 281-309.

Baddeley, A. (1996). Exploring the central executive. *Quarterly Journal Of Experimental Psychology*, **49A**(1), 5-28.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 8, pp. 47-89). New York: Academic Press.

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York: W H Freeman.

Bartlett, F. (1953). Psychological criteria of fatigue. In W. F. Floyd & A. T. Welford (Eds.), *Fatigue* (pp. 1-5). London: H K Lewis and Co.

Bartlett, F. C. (1943). Fatigue following highly skilled work. *Proceedings of the Royal Society (B)*, **131**, 147-257.

Bartley, S. H. (1965). *Fatigue: Mechanism and Management*. Springfield, Illinois: Charles C Thomas.

Bartley, S. H. (1970). The homeostatic and comfort perceptual systems. *Journal of Psychology*, **75**, 157-162.

Bartley, S. H., & Chute, E. (1947). *Fatigue and Impairment in Man*. New York: McGraw-Hill.

Beatty, J. (1982a). Phasic not tonic pupillary responses vary with auditory vigilance performance. *Psychophysiology*, **19**(2), 167-172.

Beatty, J. (1982b). Task-evoked pupillary responses, processing load and the structure of processing resources. *Psychological Bulletin*, **91**(2), 276-292.

Beatty, J. (1986). Computation, control and energetics: a biological perspective. In G. R. J. Hockey, Gaillard, A. W. K. & Coles, M. G. H. (Ed.), *Energetical states underlying task performance* (pp. 43-51). Dordrecht: Martinus Nijhoff.

Benton, D. (1987). Adrenal hormone production as indices of occupational stress. In A. Gale & B. Christie (Eds.), *Psychophysiology And The Electronic Workplace* (pp. 275-291). Chichester: Wiley.

Biers, D. W., & McInerney, P. (1988). *An alternative to measuring subjective workload: Use of SWAT without the card sort.* Paper presented at the Human Factors Society 32nd Annual Meeting, Anaheim, California.

Bills, A., G. (1931). Blocking: a new principle of mental fatigue. *American Journal of Psychology*, **43**, 230-245.

Binswanger, H. (1991). Volition as cognitive self-regulation. *Organizational Behaviour and Human Decision Processes*, **50**, 154-178.

Brichcin, M. (1982). Goal information and feedback information in the conscious control of human activity. In W. Hacker, W. Volpert, & M. Cranach (Eds.), *Cognitive and Motivational Aspects of Action* (pp. 122-133). Berlin: Huthig Verlagsgemeinschaft GmbH.

Broadbent, D. E. (1958). *Perception And Communication*. London: Pergamon Press.

Broadbent, D. E. (1963). Differences and interactions between stresses. *Quarterly Journal of Experimental Psychology*, **15**, 205-211.

Broadbent, D. E. (1971). *Decision and Stress*. London: Academic Press.

Broadbent, D. E. (1977). Levels, hierarchies and the locus of control. *The Quarterly Journal of Experimental Psychology*, **29**, 181-201.

Broadbent, D. E. (1979). Is a fatigue test now possible? *Ergonomics*, **22**(12), 1277-1290.

Brookhuis, K. A., & De Waard, D. (1993). The use of psychophysiology to assess driver status. *Ergonomics*, **36**(9), 1099-1110.

Brookhuis, K. A., De Waard, D., & Bekiaris, E. (1997). Development of a system for the detection of driver impairment. In C. Mercier-Guyon (Ed.), *Proceedings of the14th International Conference on Alcohol, Drugs and Traffic Safety* (Vol. 2, pp. 581-586). Annecy, France: CERMT.

Brookhuis, K. A., De Waard, D., & Mulder, B. (1994). Measuring driving performance by car-following in traffic. *Ergonomics*, **37**(3), 427-434.

Brown, A. L. (1988). Metacognition, executive control, self-regulation and other, even more mysterious mechanisms. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, Motivation and Understanding*. Hillsdale, NJ: Erlbaum.

Brown, I. D. (1990). Drivers' margins of safety considered as a focus for research on error. *Ergonomics*, **33**(10-11), 1307-1314.

Brown, I. D. (1994). Driver fatigue. *Human Factors*, **36**(2), 298-314.

Byers, J. C., Bittner, A. C., & Hill, S. G. (1989). Traditional and raw task load index (TLX) correlations: are paired comparisons necessary? In A. Mital (Ed.), *Advances in Industrial Ergonomics and Safety* (Vol. 1, pp. 481-488). London: Taylor and Francis.

Cameron, C. (1973). A theory of fatigue. *Ergonomics*, **16**(5), 633-648.

Cannon, W. (1932). *The Wisdom of the Body*. New York: W. W. Norton and Co.

Carver, C. S., & Scheier, M. F. (1981). *Attention and Self-Regulation: A Control Theory Approach to Human Behaviour*. New York: Springer-Verlag.

Carver, C. S., & Scheier, M. F. (1990). Origin and function of positive and negative affect: A control-process view. *Psychological Review*, **97**, 19-35.

Clayton, A. B. (1980). Effects of alcohol on driving skills. In M. Sandler (Ed.), *Psychopharmacology of Alcohol* (pp. 73-78). New York: Raven Press.

Cohen, R. M., Semple, W. E., Gross, M., Holcomb, H. H., Dowling, M. S., & Nordahl, T. E. (1988). Functional localisation of sustained attention: Comparison to sensory stimulation in the absence of instruction. *Neuropsychiatry, Neuropsychology and Behavioural Neurology*, **1**(1), 3-20.

Cox, T., Cox, S., & Thirlaway, M. (1983). The psychological and physiological response to stress. In A. Gale & J. A. Edwards (Eds.), *Physiological Correlates Of Human Behaviour: Basic Issues* (Vol. 1, pp. 254-276). London: Academic Press.

Craig, A., & Cooper, R. E. (1992). Symptoms of acute and chronic fatigue. In A. P. Smith & D. M. Jones (Eds.), *Handbook of Human Performance* (Vol. 3, pp. 289-339). London: Academic Press.

Crawford, A. (1961). Fatigue and driving. *Ergonomics*, **4**, 143-151.

Davies, D. R. (1983). Attention, arousal and effort. In A. E. Gale (Ed.), *Physiological correlates of human behaviour* (Vol. 2, pp. 9-34). London: Academic Press.

Davies, D. R., Jones, D. M., & Taylor, A. (1984). Selective- and sustained-attention tasks: individual and group differences. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of Attention* (pp. 395-447). Orlando: Academic Press.

Davies, D. R., & Parasuraman, R. (1982). *The Psychology Of Vigilance*. London: Academic Press.

De Waard, D. (1991). *Driving behaviour on a high accident rate motorway in the Netherlands.* Paper presented at the Man In Complex Systems: Proceedings of the Europe Chapter of the Human Factors Society Annual Meeting, October 1991, Lund.

De Waard, D. (1996). *The Measurement of Driver Mental Workload,* Rijksuniversiteit Groningen, Groningen, The Netherlands.

De Waard, D., & Brookhuis, K. A. (1991). Assessing driver status: a demonstration experiment on the road. *Accident Analysis and Prevention,* **23**(4), 297-301.

Derrick, W. L. (1988). Dimensions of operator workload. *Human Factors,* **30**(1), 95-110.

Desmond, P. A. (1998). *Dimensions of motivation and fatigue-related impairments: A study of simulated driving.* Paper presented at the Human Factors and Ergonomics Society Annual Meeting.

Desmond, P. A., & Matthews, G. (1997). Implications of task-induced fatigue effects for in-vehicle countermeasures to driver fatigue. *Accident Analysis and Prevention,* **29**(4), 515-524.

Desmond, P. A., Matthews, G., & Hancock, P. A. (1997). *Development and validation of the driving fatigue scale.* Paper presented at the 14th International Conference on Alcohol, Drugs and Traffic Safety, Annecy, France.

Dinges, D. F., & Kribbs, N. B. (1991). Performing while sleepy: effects of experimentally-induced sleepiness. In T. H. Monk (Ed.), *Sleep, Sleepiness and Performance* (pp. 97-128). New York: Wiley.

Donchin, E. (1984). *The use of ERPs to monitor non-conscious mentation.* Paper presented at the 20th Annual Conference on Manual Control and Mental Workload, Moffet Field, CA.

Dornic, S., Ekehammar, B., & Laaksonen, T. (1991). Tolerance for mental effort: Self-ratings related to perception, performance and personality. *Personality and Individual Differences,* **12**(3), 313-319.

Duffy, E. (1962). *Activation and Behaviour.* New York: Wiley.

Dureman, E. I., & Boden, C. (1972). Fatigue in simulated car driving. *Ergonomics,* **15**(3), 299-308.

Duval, S., & Wicklund, R. A. (1972). *A Theory of Objective Self Awareness.* New York: Academic Press.

Easterbrook, J. A. (1959). The effect of emotion on cue utilisation and the organisation of behaviour. *Psychological Review,* **66**, 183-201.

Eysenck, M. W. (1982). *Attention and Arousal.* Berlin: Springer-Verlag.

Eysenck, M. W. (1986). Individual differences in anxiety, cognition and coping. In G. R. J. Hockey, G. A. W. K., & M. G. H. Coles (Eds.), *Energetical issues in research on human information processing* (pp. 255-283). Dordrecht: Martinus Nijhoff.

Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition and Emotion,* **6**, 409-434.

Fairclough, S. H. (1997). Monitoring driver fatigue via driver performance. In Y. I. Noy (Ed.), *Ergonomics and Safety of Intelligent Driver Interfaces* (pp. 363-380). Mahwah, NJ: Erlbaum.

Fairclough, S. H., & Graham, R. (1999). Impairment of driving performance caused by sleep deprivation or alcohol: A comparative study. *Human Factors*, **41**(1), 118-128.

Fairclough, S. H., & Hirst, S. J. (1993). Driver status monitoring: Can in-vehicle IT detect the impaired driver? In E. J. Lovesey (Ed.), *Contemporary Ergonomics 1993* (pp. 469-474). London: Taylor and Francis.

Fairclough, S. H., May, A. J., & Carter, C. (1997). The effect of time headway feedback on following behaviour. *Accident Analysis and Prevention*, **29**(3), 387-397.

Fairclough, S. H., & Ward, N. J. (1996). A protocol for the assessment of subjective sleepiness. In H. F. Society (Ed.), *Human Factors and Ergonomics Society 40th Annual Meeting* (Vol. 2, pp. 1283). Philadephia: Human Factors Society.

Farmer, R., & Sundberg, N. D. (1986). Boredom Proneness - the development and correlates of a new scale. *Journal of Personality Assessment*, **50**(1), 4-17.

Fenigstein, A., Scheier, M. F., & Buss, A. H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology*, **43**(4), 522-527.

Fisk, A. D., Ackerman, P. L., & Schneider, W. (1987). Automatic and controlled processing theory and its applications to human factors problems. In P. A. Hancock (Ed.), *Human Factors Psychology* (pp. 159-197). Amsterdam: Elsevier.

Frankenhaeuser, M. (1980). Psychoneuroendocrine approaches to the study of a stressful person-environment transactions. In H. Selye (Ed.), *Selye's Guide To Stress Research* (Vol. 1, pp. 46-70). New York: Von Nostrand Reinhold.

Frankenhaeuser, M. (1987). A psychobiological framework for research on human stress and coping. In M. H. Appley & R. Trumbull (Eds.), *Dynamics Of Stress: Physiological, psychological and social perspectives* (pp. 101-116). New York: Plenum.

Frankenhaeuser, M., & Johansson, G. (1986). Stress at work: Psychobiological and psychosocial aspects. *International Review of Applied Psychology*, **36**, 287-299.

Frankenhaeuser, M., Lundberg, U., & Forsman, L. (1980). Dissociation between sympathetic-adrenal and pituitary-adrenal responses to an achievement situation characterised by high contrallability: comparison between Type A and Type B males and females. *Biological Psychology*, **10**, 79-91.

Frankenhaeuser, M., Nordheden, B., Myrsten, A. L., & Post, B. (1971). Psychophysiological reactions to understimulation and overstimulation. *Acta Psychologica*, **35**, 298-308.

Frese, M., & Zapf, D. (1994). Action as the core of work psychology: A German approach. In H. C. Triandis, M. D. Dunnette, & L. M. Hough (Eds.), *Handbook of Industrial and Organizational Psychology* (Vol. 4, pp. 271-340). Palo Alto, CA: Consulting Psychologists Press Inc.
Fuller, R. G. C. (1984). Prolonged driving in convoy: the truck driver's experience. *Accident Analysis and Prevention*, **16**(5/6), 371-382.
Gaillard, A. W. K. (1993). Comparing the concepts of mental load and stress. *Ergonomics*, **36**(9), 991-1006.
Gale, A. (1977). Some EEG correlates of sustained attention. In R. R. Mackie (Ed.), *Vigilance: theory, operational performance and physiological correlates* (pp. 263-283). New York: Plenum Press.
Gawron, V. J., & Ranney, T. A. (1988). The effects of alcohol dosing on driving performance on a closed course and in a driving simulator. *Ergonomics*, **31**(9), 1219-1244.
Godthelp, J., Milgram, P., & Blaauw, G. J. (1984). The development of a time-related measure to describe driving strategy. *Human Factors*, **26**, 257-268.
Gonzalez-Gonzalaz, J., Llorens, A. M., Novoa, A. M., & Valeriano, J. J. C. (1992). Effect of acute alcohol ingestion on short-term heart rate fluctuations. *Journal of Studies on Alcohol*, **53**(1), 86-90.
Gopher, D., & Braune, R. (1984). On the psychophysics of workload: Why bother with subjective measures? *Human Factors*, **26**(5), 519-532.
Gopher, D., & Donchin, E. (1986). Workload - an examination of the concept. In K. Boff, Kaufman, L., & Thomas, J.P. (Ed.), *Handbook of Human Perception and Performance* (Vol. 2, pp. 41.1-41.47). New York: Wiley.
Hacker, W., Plath, H. E., Richter, P., & Zimmer, K. (1978). Internal representation of task structure and mental load of work: approaches and methods of assessment. *Ergonomics*, **21**(3), 187-194.
Haier, R. J., Siegel, B. V., Nuechterlein, K. H., Hazlett, E., Wu, J. C., Paek, J., Browning, H. L., & Buchsbaum, M. S. (1988). Cortical glucose metabolic rate correlates of abstract reasoning and attention studied with positron emission tomography. *Intelligence*, **12**, 199-217.
Hamilton, P., Hockey, G. R. J., & Rejman, M. (1975). The place of the concept of activation in human information processing theory: An integrative approach. In S. Dornic (Ed.), *Attention and performance* (Vol. VI, pp. 463-486).
Hammerton, M. (1981). Tracking. In D. Holding (Ed.), *Human Skills* . Chichester: Wiley.
Hancock, P. A. (1986). Stress and adaptability. In G. R. J. Hockey, Gaillard, A. W. K. & Coles, M. G. H. (Ed.), *Energetical issues in research on human information processing* (pp. 243-250). Dordrecht: Martinus Nijhoff.

Hancock, P. A. (1988). Arousal theory, stress and performance: Problesm of incorporating energetic aspects of behavior into human-maching systems function. In L. S. Mark, J. S. Warm, & R. L. Huston (Eds.), *Ergonomics and human factors* (pp. 170-179). New York: Springer-Verlag.

Hancock, P. A. (1989). The perception of performance failure and task demand on the perception of mental workload. *Applied Ergonomics*, **20**(3), 197-205.

Hancock, P. A. (1996). Effects of control order, augmented feedback, input device and practice on tracking performance and perceived workload. *Ergonomics*, **39**(9), 1146-1162.

Hancock, P. A., & Caird, J. K. (1993). Experimental evaluation of a model of mental workload. *Human Factors*, **35**(3), 413-429.

Hancock, P. A., & Verwey, W. B. (1997). Fatigue, workload and adaptive driver systems. *Accident Analysis and Prevention*, **29**(4), 495-506.

Hancock, P. A., & Warm, J. S. (1989). A dynamic model of stress and sustained attention. *Human Factors*, **31**(5), 519-537.

Hancock, P. A., Wulf, G., Thom, D., & Fassnacht, P. (1990). Driver workload during different driving maneuvers. *Accident Analysis and Prevention*, **22**(3), 281-290.

Hart, S. G., & Staveland, L. E. (1988). Development of the NASA-TLX (Task Load Index): results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (pp. 139-183). Amsterdam: North-Holland.

Haworth, N. L., Triggs, T. J., & Grey, E. M. (1988). *Driver fatigue: concepts, measurement and crash countermeasures* (CR 72). Human Factors Group, Monash University, Australia.

Haworth, N. L., & Vulcan, P. (1991). *Testing of commercially available fatigue monitors* (15). Accident Research Centre, Monash University, Australia.

Heatherton, T. F., & Polivy, J. (1991). Development of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology*, **60**, 895-910.

Hendy, K. C., Liao, J., & Milgram, P. (1997). Combining time and intensity effects in assessing operator information-processing load. *Human Factors*, **39**(1), 30-47.

Hockey, G. R. J. (1986a). Changes in operator efficiency as a function of environmental stress, fatigue, and circadian rhythms. In K. R. Boff, I. Kaufman, & J. P. Thomas (Eds.), *Handbook of Perception and Human Performance* (Vol. 2, pp. 44.3-44.44). New York: Wiley.

Hockey, G. R. J. (1986b). A state control theory of adaptation and individual differences in stress management. In G. R. J. Hockey, Gaillard, A. W. K. & Coles, M. G. H. (Ed.), *Energetical issues in research on human information processing* (pp. 285-298). Dordrecht, The Netherlands: Martinus Nijhoff.

Iockey, G. R. J. (1993). Cognitive-energetical control mechanisms in the management of work lemands and psychological health. In A. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, 4wareness and Control.* (pp. 328-345). Oxford: Clarendon Press.

Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: a cognitive-energetical framework. *Biological Psychology*, **45**, 73-93.

Hockey, G. R. J., Coles, M. G. H., & Gaillard, A. W. K. (1986). Energetical issues in research on human information processing. In G. R. J. Hockey, A. W. K. Gaillard, & M. G. H. Coles (Eds.), *Energetics and Human Information Processing* (pp. 3-21). Dordrecht: Martinus Nijhoff.

Hockey, G. R. J., & Hamilton, P. (1983). The cognitive patterning of stress states. In G. R. J. Hockey (Ed.), *Stress and human fatigue in human performance* (pp. 331-362). Chichester: John Wiley and Sons.

Ingram, R. E. (1990). Self-focused attention in clinical disorders: A review and conceptual model. *Psychological Bulletin*, **107**(2), 156-176.

James, W. (1890). *The Principles of Psychology.* New York: Holt.

Johns, M. W. (1992). A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. *Sleep*, **14**, 540-545.

Jorna, P. G. A. M. (1992). Spectral analysis of heart rate and psychological state: a review of its validity as a workload index. *Biological Psychology*, **34**, 237-257.

Kahneman, D. (1970). Remarks on attentional control. In A. F. Sanders (Ed.), *Attention and Performance* (Vol. 3, pp. 118-131). Amsterdam: North-Holland.

Kahneman, D. (1973). *Attention and Effort.* Englewood Cliffs N. J.: Prentice-Hall.

Karasek, R. A. (1979). A stress management model of job strain. *Administrative Science Quarterly*, **24**, 285-308.

Kecklund, G., & Akerstedt, T. (1993). Sleepiness in long distance truck driving: an ambulatory EEG study of night driving. *Ergonomics*, **36**(9), 1007-1018.

Kimble, G. A., & Perlmuter, L. C. (1970). The problem of volition. *Psychological Review*, **77**(5), 361-384.

Kjellberg, A. (1977a). Sleep deprivation and some aspects of performance II. Lapses and other attentional effects. *Waking and Sleeping*, **1**, 145-148.

Kjellberg, A. (1977b). Sleep deprivation, arousal and performance. In R. R. Mackie (Ed.), *Vigilance: theory, operational performance and physiological correlates* (pp. 529-535). New York: Plenum Press.

Kramer, A., & Spinks, J. (1991). Capacity views of human information processing. In J. R. Jennings, & Coles, M. G. H. (Ed.), *Handbook of Cognitive Psychophysiology: Central and Nervous Systems Approaches* (pp. 179-249). New York: Wiley.

Lacey, J. I. (1967). Somatic response patterning and stress: Some revisions of activation theory. In M. H. Appley & R. Trumbull (Eds.), *Psychological Stress: Issues In Research* (pp. 14-42). New York: Appleton-Century-Crofts.

Langar, E. J., & Imber, L. G. (1979). When practice makes imperfect: The debilitating effects of overlearning. *Journal of Personality and Social Psychology*, **37**(11), 2014-2024.

Lazarus, R. S., & Folkman, S. (1984). *Stress, Appraisal and Coping*. New York: Springer-Verlag.

Locke, E. A., & Latham, G. P. (1990). *A Theory of Goal Setting and Task Performance*. Englewood Cliffs, NJ: Prentice-Hall.

Logan, G. D. (1985). Executive control of thought and action. *Acta Psychologica*, **60**, 193-210.

Lord, R. G., & Levy, P. E. (1994). Moving from cognition to action: A control theory perspective. *Applied Psychology: An International Review*, **43**, 335-366.

Lundberg, U., & Forsman, L. (1979). Adrenal-medullary and adrenal-cortical responses to understimulation and overstimulation: comparison between Type A and Type B persons. *Biological Psychology*, **9**, 79-89.

Lundberg, U., & Frankenhaeuser, M. (1978). Psychophysiological reactions to noise as modified by personal control over noise intensity. *Biological Psychology*, **6**, 51-59.

Mackie, R. R., & Wylie, C. D. (Eds.). (1990). *Countermeasures to loss of alertness in truck drivers: Theoretical and practical consequences* . Bron, France.: Paradigme.

Madsen, P. L. (1993). Blood flow and oxygen uptake in the human brain during various states of sleep and wakefulness. *Acta Neurologica Scandinavica*, **88**, S5-27.

Malmo, R. B. (1959). Activation: a neurophysiological dimension. *Psychological Review*, **66**, 367-386.

Mandler, G., Mandler, J. M., & Uviller, E. T. (1958). Autonomic feedback: The perception of autonomic activity. *Journal of Abnormal and Social Psychology*, **56**, 367-373.

Mascord, D. J., & Heath, R. A. (1992). Behavioural and physiological indices of fatigue in a visual tracking task. *Journal of Safety Research*, **23**, 19-25.

Matthews, G. (1992). Mood. In A. P. Smith & D. M. Jones (Eds.), *Handbook of Human Performance* (Vol. 3, pp. 161-193). London: Academic Press.

Matthews, G., Davies, D. R., & Holley, P. J. (1993). Cognitive predictors of vigilance. *Human Factors*, **35**(1), 3-24.

Matthews, G., & Desmond, P. A. (1998). Personality and the multiple dimensions of task-induced fatigue: a study of simulated driving. *Personality and Individual Differences*, **25**, 443-458.
Matthews, G., Jones, D. M., & Chamberlain, A. G. (1990). Refining the measurement of mood: The UWIST Mood Adjective Checklist. *British Journal of Psychology*, **81**, 17-42.
Matthews, G., Joyner, L., Gilliland, K., Campbell, S., Falconer, S., & Huggins, J. (1997). Validation of a comprehensive stress state questionnaire: Towards a state 'Big Three'? In I. Mervielde, I. J. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe* (Vol. 7,). Tilburg: Tilburg University Press.
Matthews, G., Schwean, V. L., Campbell, S. E., Saklofske, D. H., & Mohamed, A. A. R. (in press). Personality, self-regulation and adaptation: A cognitive-social framework. In M. Boekarts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-regulation* . New York: Academic Press.
Matthews, G., Sparkes, T. J., & Bygrave, H. M. (1996). Attentional overload, stress and simulated driving performance. *Human Performance*, **9**(1), 77-101.
Matthews, K. A., Carver, C. S., & Scheier, M. F. (1982). Focus of attention and awareness of bodily sensations. In G. Underwood (Ed.), *Aspects Of Consciousness* (Vol. 3, pp. 165-197). London: Academic Press.
Maycock, G. (1995). *Driver sleepiness as a factor in car and HGV accidents* (Report 169). Transport Research Laboratory, Crowthorne, UK.
Mayleben, D. W., Warm, J. S., Dember, W. N., Rosa, R. R., Shear, P., Temple, J., & Parasuraman, R. (1998). *Cerebral blood flow during sustained attention*. Paper presented at the Third Automation and Human Performance Conference, Norfolk, Virginia.
McDonald, N. (1987). Fatigue and driving. *Alcohol, Drugs and Driving*, **5**(3), 185-192.
McLean, J. R., & Hoffman, E. R. (1975). Steering reversals as a measure of driving performance and steering task difficulty. *Human Factors*, **17**, 248-256.
Meijman, T. F. (1995). *Mental fatigue and the temporal structuring of working times*. Paper presented at the 39th Annual Meeting of the Human Factors and Ergonomics Society, San Diego, California.
Meijman, T. F. (1997). Mental fatigue. *International Journal of Industrial Ergonomics*, **20**(1), 31-38.
Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the Structure of Behaviour*. New York.
Moray, N. (1967). Where is capacity limited? A survey and a model. *Acta Psychologica*, **27**, 84-92.
Moray, N. (1982). Subjective mental workload. *Human Factors*, **24**(1), 25-40.
Moray, N., King, I., Turksen, B., & Waterton, K. (1987). A closed loop causal model of workload based on a comparison of crisp and fuzzy measurement techniques. *Human Factors*, **29**, 339-348.

Muckler, F. A., & Seven, S. A. (1992). Selecting performance measures: "objective" versus "subjective" measurement. *Human Factors*, **34**(4), 441-455.
Mulder, G. (1979a). Mental load, mental effort and attention. In N. Moray (Ed.), *Mental Workload: Its Theory and Measurement* (pp. 327-343). New York: Plenum Press.
Mulder, G. (1979b). Sinus arrhymia and mental workload. In N. Moray (Ed.), *Mental Workload: Its Theory and Measurement* (pp. 327-343). New York: Plenum Press.
Mulder, G. (1985). Attention, effort and sinus arrhymia: how far are we? In J. F. Orlebeke, G. Mulder, & L. J. P. Van Doornen (Eds.), *Psychophysiology of Cardiovascular Control: Models, methods and data.* (pp. 407-423). New York: Plenum Press.
Mulder, G. (1986). The concept and measurement of mental effort. In G. R. J. Hockey, Gaillard, A. W. K. & Coles, M. G. H. (Ed.), *Energetical issues in research on human information processing* (pp. 175-198). Dordrecht, The Netherlands: Martinus Nijhoff.
Mulder, G., & Mulder, L. M. J. (1981). Information processing and cardiovascular control. *Psychophysiology*, **8**(4), 392-402.
Mulder, G., & Van Der Meulen, W. R. E. H. (1973). Mental load and the measurement of heart rate variability. *Ergonomics*, **16**(1), 69-83.
Mulder, L. J. M., & Schweizer, D. A. (1993). *Carspan: Cardiovascular Experiments Analysis Environment* . iec ProGamma, Groningen, The Netherlands.
Mulders, H. P. G., Meijman, T. F., O'Hanlon, J. F., & Mulder, G. (1982). Differential psychophysiological reactivity of city bus drivers. *Ergonomics*, **25**(11), 1003-1011.
Muscio, B. (1921). Is a fatigue test possible? *British Journal of Psychology*, **12**, 31-46.
Naatanen, R., & Summala, H. (1975). *Effort in driving as a traffic-safety problem*. Paper presented at the Physical work and effort: Proceedings of the first annual symposium, Stockholm.
Navon, D. (1984). Resources - a theoretical soupstone? *Psychological Review*, **91**, 216-234.
Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review*, **86**(3), 214-225.
Nelson, T. M. (1987). Subjective factors related to fatigue. *Alcohol, Drugs and Driving*, **5**(3), 193-214.
Nilsson, T., Nelson, T. M., & Carlson, D. (1997). Development of fatigue symptoms during simulated driving. *Accident Analysis and Prevention*, **29**(4), 479-488.
Norman, D. A. (1981). Categorization of action slips. *Psychological Review*, **88**(1), 1-16.
Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, **7**, 44-64.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and Self-Regulation* (Vol. 4, pp. 1-18). New York: Plenum Press.

Nowlis, V. (1965). Research with the mood adjective checklist. In S. S. Tomkins & C. E. Izard (Eds.), *Affect, Cognition and Personality* (pp. 352-389). New York: Springer.

O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology. In K. Boff, Kaufman, L., & Thomas, J.P. (Ed.), *Handbook of Human Perception and Performance* (Vol. 2, pp. 42.1-42.49). New York: Wiley.

O'Hanlon, J. F., & Beatty, J. (1977). Concurrence of electroencephalographic and performance changes during a simulated radar watch and some implications for the arousal theory of vigilance. In R. R. Mackie (Ed.), *Vigilance: theory, operational performance and physiological correlates* (pp. 189-201). New York: Plenum Press.

Paoli, P. (1997). *Second European Survey on the Work Environment 1996-1998* . European Foundation for the Improvement of Living and Working Conditions.

Parasuraman, R. (1983). Vigilance, arousal and the brain. In A. Gale & J. A. Edwards (Eds.), *Physiological correlates of human behaviour* (Vol. 2, pp. 35-55). London: Academic Press.

Parasuraman, R. (1986). Vigilance, monitoring and search. In K. Boff, Kaufman, L., & Thomas, J.P. (Ed.), *Handbook of Human Perception and Performance* (Vol. 2, pp. 43.1-43.39). New York: Wiley.

Parkes, A. M., Fairclough, S. H., & Ross, T. (1991). Multi-level evaluations of in-vehicle route navigation system. In E. J. Lovesey (Ed.), *Contemporary Ergonomics 1991* (pp. 157-162). London: Taylor and Francis.

Pashler, H. E. (1998). *The Psychology Of Attention*. Mass: MIT Press.

Paus, T., Mates, J., Radil, T., Hampl, R., & Husek, P. (1988). Controlled and automatic processing modes in vigilance situation: Neuroendocrine correlates of effort. *Activas-Nervosa-Superior*, **30**(4), 289-290.

Pawlik, G., & Heiss, W. D. (1989). Positron emission topography and neuropsychological function. In E. D. Bigler, R. A. Yeo, & E. Turkheimer (Eds.), *Neuropsychological Function and Brain Imaging. Critical Issues in Neuropsychology* (pp. 65-138). New York: Plenum.

Pennebaker, J. W., & Skelton, J. A. (1978). Psychological parameters of physical symptoms. *Personality and Social Psychology Bulletin*, 4(4), 524-530.

Pribram, K. H. (1980). Mind, brain and consciousness: The organisation of competence and conduct. In J. M. Davidson & R. J. Davidson (Eds.), *Psychobiology of Consciousness* (pp. 47-63). New York: Plenum.

Rabbitt, P. M. A. (1978). Detection of errors by skilled typists. *Ergonomics*, **21**, 945-958.
Rabbitt, P. M. A. (1981). Sequential reactions. In D. H. Holding (Ed.), *Human Skills*. New York: Wiley.
Rabbitt, P. M. A., Cumming, G., & Vyas, S. V. (1978). Some errors of perceptual analysis in visual search can be detected and corrected. *Quarterly Journal of Experimental Psychology*, **30**, 319-332.
Rasmussen, J., & Jensen, A. (1974). Mental procedures in real-life tasks: a case study of electronic troubleshooting. *Ergonomics*, **17**, 293-307.
Reason, J. (1987). Generic Error-Modelling System (GEMS): A cognitive framework for locating common human error forms. In J. Rasmussen, K. Duncan, & J. Leplat (Eds.), *New Technology and Human Error* (pp. 63-83). Chichester: Wiley.
Reason, J. (1990). *Human Error*. Cambridge: Cambridge University Press.
Reivich, M., Alavi, A., Gur, R. C., & Greenberg, J. (1985). Determination of local cerebral glucose metabolism in humans: Methodology and applications to the study of sensory and cognitive stimuli. In L. Sokoloff (Ed.), *Brain Imaging and Brain Function* (pp. 105-119). New York: Raven Press.
Riemersma, J. B. J., & Biesta, P. W. (1978). *On the effects of prolonged night driving*. Paper presented at the Driver Fatigue in Road Traffic Accidents: contributions to workshops on physiological, psychological and sociological aspects of the problem.
Riemersma, J. B. J., Sanders, A. F., Wildervanck, C., & Gaillard, A. W. K. (1977). Performance decrement during prolonged night driving. In R. R. Mackie (Ed.), *Vigilance: theory, operational performance and physiological correlates* (pp. 41-58). New York: Plenum Press.
Ryan, T. A. (1953). Muscular potentials as indicators of effort in visual tasks. In W. F. Floyd & A. T. Welford (Eds.), *Fatigue* (pp. 109-116). London: H K Lewis and Co.
Sanders, A. F. (1983). Towards a model of stress and human performance. *Acta Psychologica*, **53**, 61-97.
Sanders, A. F. (1985). Drugs, driving and the measurement of human performance. In De Gier & J. F. O'Hanlon (Eds.), *Drugs and Driving* (pp. 3-16): Taylor and Francis.
Sarason, I. G., Sarason, B. R., Keefe, D. E., Hayes, B. E., & Shearin, E. N. (1986). Cognitive interference: Situational determinants and traitlike characteristics. *Journal of Personality and Social Psychology*, **57**, 691-706.
Sarason, I. G., Sarason, B. R., & Pierce, G. R. (1990). Anxiety, cognitive interference and performance. *Journal of Social Behaviour and Personality*, **5**, 1-18.

Sarason, I. G., Sarason, B. R., & Pierce, G. R. (1995). Cognitive interference: At the intelligence-personality crossroads. In D. H. Saklofske & M. Zeidner (Eds.), *International Handbook of Personality and Intelligence* . New York: Plenum.

Sawin, D. A., & Scerbo, M. W. (1995). Effects of instruction type and boredom proneness in vigilance: implications for boredom and workload. *Human Factors*, **37**(4), 752-765.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing I. Detection, search and attention. *Psychological Review*, **84**(1), 1-66.

Schonpflug, W. (1983). Coping efficiency and situational demands. In G. R. J. Hockey (Ed.), *Stress and Fatigue in Human Performance* (pp. 299-330). Chichester: Wiley.

Schonpflug, W. (1985). Goal-directed behaviour as a source of stress: Psychological origins and consequences of inefficiency. In M. Frese & J. Sabini (Eds.), *Goal-Directed Behaviour: The Concept of Action in Psychology* (pp. 172-188). Hillsdale, NJ: Erlbaum.

Schonpflug, W. (1986a). Behaviour economics as an approach to stress theory. In M. H. Appley & R. Trumbull (Eds.), *Dynamics of Stress: Physiological, Psychological and Social Perspectives* (pp. 81-98). New York: Plenum.

Schonpflug, W. (1986b). Effort regulation and individual differences in effort expenditure. In G. R. J. Hockey, A. W. K. Gaillard, & M. G. H. Coles (Eds.), *Energetics and Human Information Processing* (pp. 271-283). Dordrecht: Martinus Nijhoff.

Schultz, P., & Schonpflug, W. (1982). Regulatory activities during states of stress. In H. W. Krohne & L. Laux (Eds.), *Achievement, Stress and Anxiety* (pp. 51-73). Washington: Hemisphere.

Shallice, T. (1978). The dominant action system: An information-processing approach to consciousness. In K. S. Pape & J. L. Singer (Eds.), *The Stream Of Consciousness* (Vol. lib153,). New York: Plenum.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society London*, **B**(298), 199-209.

Shallice, T. (1988). *From Neuropsychology to Mental Structure*. Cambridge: Cambridge University Press.

Shallice, T., & Burgess, P. (1993). Supervisory control of action and thought selection. In A. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection , Awareness and Control* (pp. 171-187). Oxford: Clarendon Press.

Sheridan, T. B., Meyer, J. E., Roy, S. H., Decker, K. S., Yanagishma, T., & Kishi, Y. (1991). Physiological and psychological evaluations of driver fatigue during long-term driving. *SAE Technical Paper Series*, **910116**.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, **84**(2), 127-190.
Sperandio, J.-C. (1978). The regulation of working methods as a function of workload among Air Traffic Controllers. *Ergonomics*, **21**(3), 195-202.
Stassen, H. G. (1989). Internal representation, internal model, human performance model and mental workload. In J. Ranta (Ed.), *Analysis, Design and Evaluation of Man-Machine Systems* (pp. 23-32). Oxford: Pergamon.
Strelau, J. (1985). Temperament and personality: Pavlov and beyond. In J. Strelau, F. Farley, & A. Gales (Eds.), *The Biological Bases of Personality and Behaviour* (Vol. 1, pp. 25-43). New York: Hemisphere.
Tattersall, A. J., & Hockey, G. R. J. (1995). Level of operator control and changes in heart rate variability during simulated flight maintenance. *Human Factors*, **37**(4), 682-698.
Teigen, K. H. (1994). Yerkes-Dodson: A law for all seasons. *Theory and Psychology*, **4**(4), 525-547.
Thackray, R. I., Bailey, J. P., & Touchstone, R. M. (1977). Physiological, subjective and performance correlates of reported boredom and monotony while performing a simulated radar control task. In R. R. Mackie (Ed.), *Vigilance: theory, operational performance and physiological correlates* (pp. 203-215). New York: Plenum.
Thayer, R. E. (1967). Measurement of activation through self-report. *Psychological Reports*, **20**, 663-678.
Thayer, R. E. (1989). *The Biopsychology Of Mood And Arousal*. New York: Oxford University Press.
Thierry, H., & Meijman, T. (1994). Time and behaviour at work. In H. C. Triandis, M. D. Dunnette, & L. M. Hough (Eds.), *Handbook of Industrial and Organizational Psychology* (Vol. 4, pp. 341-414). Palo Alto, CA: Consulting Psychologists Press Inc.
Thorndike, E. (1900). Mental fatigue. *Psychology Review*, **7**, 466-482.
Tsaneva, N., & Markov, S. (1971). A model of fatigue. *Ergonomics*, **14**(1), 11-16.
Tulga, M. K., & Sheridan, T. B. (1980). Dynamic decision and workload in multitask supervisory control. *IEEE Transaction on Systems, Man and Cybernetics*, **SMC-10**, 217-232.
Ullsperger, P., Metz, A., & Gille, H. (1988). The P300 component of the event-related brain potential and mental effort. *Ergonomics*, **31**, 1127-1137.
Van Dellen, H. J., Aasman, J., Mulder, L. J. M., & Mulder, G. (1985). Time domain versus frequency domain measures of heart rate variability. In J. F. Orlebeke, G. Mulder, & L. J. P. Van Doornen (Eds.), *Psychophysiology of Cardiovascular Control: Models, methods and data.* (pp. 353-374). New York: Plenum Press.

Van Toller, C. (1983). Biochemistry of the nervous system. In A. Gale & J. A. Edwards (Eds.), *Physiological Correlates of Human Behaviour* (Vol. 1, pp. 79-102).

Van Winsum, W. (1996). *From Adaptive Control to Adaptive Driver Behaviour*, Rijksuniversiteit Groningen, The Netherlands.

Vasey, M. W., & Thayer, J. F. (1987). The continuing problem of false positives in repeated measures ANOVA in psychophysiology: a multivariate solution. *Psychophysiology*, **24**(4), 479-486.

Venables, P. H. (1984). Arousal: An examination of its status as a concept. In M. G. H. Coles, J. R. Jennings, & J. P. Stern (Eds.), *Psychological Perspectives* . New York: Van Nostrand.

Vicente, K., Thornton, D. C., & Moray, N. (1987). Spectral analysis of sinus arrhymia: a measure of mental effort. *Human Factors*, **29**(2), 171-182.

Walls, H. J., & Brownlie, A. R. (1985). *Drink, Drugs & Driving*. London: Sweet & Maxwell.

Warm, J. S., & Dember, W. N. (1998). Tests of vigilance taxonomy. In R. R. Hoffman, M. F. Sherrick, & J. S. Warm (Eds.), *Viewing Psychology As A Whole: The Integrative Science of William N Dember* (pp. 89-112). Washington D C: American Psychological Association.

Weimann, J. (1989). Mental effort in problem-solving as manifested in the power spectra of heart-interbeats intervals. In K. Landau & W. Rohmert (Eds.), *Recent Developments In Job Analysis* (pp. 227-236). London: Taylor and Francis.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, **92**, 548-573.

Wells, A. (1994). Attention and the control of worry. In G. L. C. Davey & F. Tallis (Eds.), *Worrying: Perspectives on theory, assessment and treatment* (pp. 91-114). Chichester, UK: Wiley.

Wells, A., & Matthews, G. (1994). *Attention and Emotion: A Clinical Perspective*. Hove: Erlbaum.

Wesnes, K., & Warbuton, D. (1983). Stress and drugs. In G. R. J. Hockey (Ed.), *Stress and Fatigue in Human Performance* (pp. 203-243). Chichester: Wiley.

Wickens, C. D. (1980). The structure of processing resources. In R. Nickerson, & Pew, R. (Ed.), *Attention and Performance VIII* (pp. 239-257). Hillsdale, N.J.: Erlbaum.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman, & Davies, D. R. (Ed.), *Varieties of Attention* (pp. 63-101). London: Academic Press.

Wickens, C. D. (1986a). The effect of control dynamics on performance. In K. R. Boff, I. Kaufman, & J. P. Thomas (Eds.), *Handbook of Perception and Human Performance: Cognitive Processes and Performance* (Vol. 2, pp. 39.1-39.60). New York: Wiley.

Wickens, C. D. (1986b). Gain and energetics in information processing. In G. R. J. Hockey, Gaillard, A. W. K. & Coles, M. G. H. (Ed.), *Energetical issues in research on human information processing* (pp. 373-389). Dordrecht: Martinus Nijhoff.

Wickens, C. D. (1991). Processing resources and attention. In D. L. Damos (Ed.), *Multiple-Task Performance* (pp. 3-34). London: Taylor and Francis.

Wiener, N. (1948). *Cybernetics: Control and communication in the animal and the machine.* (2 ed.). Cambridge, Mass.: M.I.T. Press.

Wierwille, W. W. (1994). *Overview of research on driver drowsiness definition and driver drowsiness detection*. Paper presented at the 14th International Conference on Enhanced Safety of Vehicles, Munich, Germany.

Wierwille, W. W., Wreggit, S. S., & Knipling, R. R. (1994). *Development of improved algorithms for on-line detection of driver drowsiness.* Paper presented at the Convergence'94, Detroit, Mich.

Wilkinson, R. T. (1962). Muscle tension during mental work under sleep deprivation. *Journal of Experimental Psychology*, **64**, 565-571.

Yeh, Y. Y., & Wickens, C. D. (1988). Dissociation of performance and subjective measures of workload. *Human Factors*, **30**(1), 111-120.

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

Yoshitake, H. (1971). Relations between the symptoms and feelings of fatigue. *Ergonomics*, **14**, 175-185.

Yoshitake, H. (1978). Three characteristic patterns of subjective fatigue symptoms. *Ergonomics*, **21**, 231-233.

Zijlstra, F. R. H., & van Doorn, L. (1985). *The construction of a scale to measure perceived effort* (P85 Z11 INT). Department of Philosophy and Social Sciences, Delft University of Technology.