THE UNIVERSITY OF HULL

The construct of psychological fatigue:
A psychometric and experimental analysis

Being a Thesis submitted for the Degree of
Doctor of Philosophy in the University of Hull

By

Fiona Earle (BSc)

May 2004

CONTENTS

170 Chapter 6: Task Demands, Workload and Subjective State

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Fatigue is a familiar and commonplace occurrence, but attempts to investigate the nature of fatigue have been inconclusive. Following more than a hundred years of extensive research, the construct is still ill-defined. This has resulted in a series of different strands of research, producing results concomitant with each researcher's own idea of what constitutes fatigue.

Two central questions remain unresolved: (1) what sort of a construct is fatigue? and (2) should fatigue be conceptualised as a single, one dimensional state, generated by a range of different conditions, or a multidimensional state, incorporating a number of distinct but related states? There is an implicit assumption within the literature (and every-day language) that there is more than one 'type' of fatigue. However, there is currently no theoretical model which outlines the types of fatigue which should be incorporated in a theoretical framework and which explains the relationships between these fatigue types.

The work presented in this thesis represents an attempt to address these issues using both psychometric and experimental approaches. Preliminary work investigated the psychometric basis for a unitary or multidimensional construct. This separately addressed the constructs of state and trait fatigue and, on the basis of the findings, state and trait multiple fatigue questionnaires were developed.

A series of four experiments were then carried out which manipulated different types of work to facilitate an investigation of the dynamic development of fatigue. The first three experiments focused on the separate effects of mental and physical fatigue, and the final experiment considered the nature of their interaction.

Both experimental and psychometric analyses supported the proposition of a multidimensional construct. The evidence in support of a multidimensional construct of trait fatigue was particularly strong. However, while the evidence in support of a multidimensional construct of state fatigue was less convincing, the experimental manipulations of different types of workload did produce states of fatigue that were subjectively different and also different patterns of fatigue after-effects.

ACKNOWLEDGEMENTS

My first thanks must go to my three supervisors who contributed to this process in very different ways. To my original supervisor, Professor Bob Hockey, I am grateful for starting me on this journey and teaching me about experimental design and statistics. To Dr Martin Crawshaw I will be eternally grateful for his enormous support and help over the past two years, particularly with regard to the preparation of the thesis (and, of course, educating me on the mysteries of the split infinitive). Finally, I would like to thank Dr Peter Clough, who has been a supervisor, a mentor and a true friend – thank you.

There are many other people in the Psychology Department to whom I am also grateful. In particular, thanks go to Dr Catherine Fritz, Jane Aisbitt and Dave Marchant, for providing helpful comments on sections of the thesis; to Dr Rowan Myron and Linda Randerson for their continuous support; and special thanks to Dr Jo Horne for helping in so many ways.

I would also like to thank Dr Dave Wastell and Martin Schlaf for writing the computer software programmes which were a fundamental part of the experimental work. To Martin I am also grateful for the numerous occasions on which he rescued me from my various computer disasters.

Much of the work presented in the thesis required a huge investment of time and effort on the part of the participants. So, thanks must also go to the students who took part in this research, as it would not have been possible without them.

Finally, I am grateful to my family and friends for the love and support they have given me throughout this process. Most of all, I am grateful to my husband Keith who has been by my side every step of the way, through the ups and the downs, and without whom I would never have got this far.

Part 1:

General Introduction and Research Background

Chapter 1

Fatigue research:

Methodological approaches and models

1.1 Summary

Fatigue research has a long history and the problems associated with this work are well documented. Essentially, there remains no agreed definition of what fatigue actually *is* and two central questions have gone unresolved: (1) what sort of a construct is fatigue? and (2) should fatigue be conceptualised as a single, one dimensional state, generated by a range of different conditions, or a multidimensional state, incorporating a number of distinct but related states?

Implicit within the literature (and every-day language) is the assumption that there is more than one 'type' of fatigue - It is not uncommon to see distinctions between *mental, physical, emotional* and *sleep-related* fatigue. However, even if we accept the proposition that there are a number of different types of fatigue, there is currently no theoretical model which outlines the types of fatigue that should be incorporated in a theoretical framework and which explains the relationships between these fatigue types.

This chapter considers the fatigue literature from the perspective of the different measurement approaches that have been adopted. A range of performance-based and subjective approaches are considered and the conceptual implications of each approach are discussed.

1.2 Introduction

In 1921 the British Journal of Psychology published a report to the Industrial Fatigue Research Board entitled "Is a fatigue test possible?". In this insightful monograph, Muscio comprehensively outlined the fundamental issues associated with the reliable measurement of fatigue. In essence, these issues centred on the problems of definition and the search for a 'characteristic expression of fatigue' - a discussion which provided strong direction for future research. However, it was Muscio's brief conclusion which has become the most memorable, and frequently quoted, maxim of the fatigue literature: Such were the difficulties of finding an adequate basis for defining fatigue, he recommended that "the term fatigue be absolutely banished from precise scientific discussion, and consequently attempts to obtain a fatigue test abandoned" (p45).

Despite this pessimistic plea, the term *fatigue* has persisted within the psychology (and physiology) literature. But the ongoing tendency to quote Muscio's conclusion is testament to the continued problems surrounding the fatigue construct: Essentially, the problems of definition persist, and there remains no agreed definition of what fatigue actually *is*. While there are some broadly common elements, such as the presence of subjective tiredness and impairments in work performance, researchers have differentially emphasised these (either implicitly or explicitly), which has inevitably resulted in a broad range of different conceptualisations of fatigue all addressed by a common term.

Inextricably bound up with the problem of defining the *nature* of fatigue definition, is the issue of *breadth* of definition: What is the relationship between the different types of fatigue? Research efforts within psychology have largely focused on the state which develops as a consequence of engaging in mental work (*mental fatigue*). However, as identified by Hockey and Meijman (1998) there is an implicit assumption within the literature that there is more than one 'type' of fatigue. It is not uncommon to see distinctions between a range of other types of fatigue, including *physical*, *emotional* and *sleep-related* fatigue. The question is whether fatigue should be considered a single, unidimensional state which is generated by a range of different conditions, or whether fatigue would be better conceptualised as multidimensional, incorporating a number of distinct but related states.

Complicating this further is the lack of agreement on how many different 'types' of fatigue there are. Even if it is assumed that there are a number of different fatigue processes, there is still no consensus on which types of fatigue should be incorporated into a conceptual framework. Typically, the types of fatigue considered depends on specific research aims, for example, research into the effects of shiftwork may consider 'sleep-related fatigue', whereas research into fatigue in a caring profession may focus on 'emotional fatigue'. Although such approaches may be justifiable in the context of individual research programmes, this has done little to contribute to a comprehensive fatigue model which explains the relationships between the fatigue types.

Such conceptual issues are not typically addressed within the literature. However, the question of dimensionality is important in both theoretical and practical terms: In theoretical terms, it is difficult to conceive of an adequate model of fatigue which does not address the issue of dimensionality. Furthermore, this issue has important practical implications with regard to the joint effects of the different sources of fatigue, as the majority of modern working environments impose different demands which may give simultaneous rise to more than one type of fatigue.

Therefore, the work presented in this thesis represents a preliminary attempt to address the issue of the dimensionality of the fatigue construct.

1.3 Approaches to fatigue measurement

While comprehensive and detailed descriptions of fatigue research have been provided elsewhere (most notably: Bartley and Chute, 1947; Cameron, 1973; Craig & Cooper, 1992) an overview of the important milestones in the literature is obviously still relevant here. However, rather than providing a chronological review of the literature, the following Chapter will consider fatigue research from the perspective of the different measurement approaches which have been employed. This is a pertinent approach for two reasons. Firstly, the important themes of research upon which this thesis is based did not develop sequentially, but could be better described as a series of parallel strands - thus making a chronological overview problematic. Secondly, because much of the research has been carried out without any attempt to define fatigue explicitly, one must look to the measurement approaches employed to gain some understanding of the implicit definitions of the construct.

1.3.1 Simple performance decrements: an analogy to physical fatigue

The onset of formal fatigue research is typically accredited to Mosso, an Italian biologist who, in 1890, published his research on muscular fatigue in man. This work demonstrated a reliable pattern of performance impairment within a muscle or group of muscles following repeated stimulation. His particularly successful approach was based on a finger ergograph, which enabled him to record this objective measure of impairment on a performance curve. This deterioration in muscle capability became known as muscular fatigue, and developed into an important area of research for physiologists. But, despite his biological focus, Mosso went beyond the concepts of muscular fatigue and distinguished between mental and physical fatigue, even noting some enduring characteristics of mental fatigue such as changes in the level of attention, and increased irritability. Furthermore, he noted that individuals differ greatly in the way they show fatigue - an idea which has received little attention in the psychological fatigue literature, but will be returned to in detail in Chapters Three and Four. However, despite these interesting observations, the enduring legacy of this work was the widespread adoption of a similar methodology to the study of mental or psychological fatigue.[1] Thus, by analogy, it was assumed that mental work would produce a similar curve of performance impairment. However, as noted by Muscio (1921b), while muscular fatigue can be reliably demonstrated, evidence of a reduced output in repetitive mental work was much less forthcoming: In an attempt to recreate a reliable performance curve, which could be taken to represent mental fatigue, researchers required participants to carry out long series of exactly repeated activity. An example of this method was the work of Rivers (1896) who employed Kraeplin's Reckoning Test. This task required participants to add, subtract, multiply or divide simple rows of digits for prolonged periods, with the aim of investigating fatigue and performance impairment. Similar was the work of McDougall (1905), which utilised a dotting machine requiring participants to make a repetitive response to a continuously changing display. While there were some benefits to such an approach, in that it advanced our understanding of diurnal performance rhythms, this was not a fruitful method for the study of task-related mental fatigue. In particular, performance did not deteriorate in a reliable curve of impairment as initially anticipated. For example, Poffenberger (1928) required participants to carry out a series of simple mental tasks for approximately 5½ hours and, in some instances, performance continued to improve

---

[1] The terms 'psychological fatigue' and 'mental fatigue' are frequently used interchangeably within the literature. As the current thesis debates the nature of the fatigue construct, it is difficult to define the terms clearly at this stage. However, mental fatigue is considered to be a more specific term which refers to a state which develops as a consequence of engaging in cognitive or mental work, whereas psychological fatigue is taken to represent a broader concept, incorporating the effects of sleep deprivation, emotional strain, etc.

throughout the experimental session. Such findings are far from unusual, as reports of almost heroic performance are littered throughout the history of fatigue research. Of further note, however, was Poffenberger's observation that there appeared to be no relationship between average loss in performance output and change in 'feelings'.

Therefore, not only did simple performance measurements show a surprising resistance to performance impairment over time, but decrements appeared to bear no relationship to changes in subjective state. These two issues have long been recognised as the central problems in fatigue research (Muscio, 1921a; Bartlett, 1943; Cameron, 1969), which has led to the development of two broad strands of research - one emphasising alternative approaches to performance impairment, focusing on either qualitative changes during complex tasks or performance after effects; the second emphasising subjective state and placing the feelings of fatigue at the centre of the construct.

However, before going on to discuss these two strands of research, and their associated conceptualisations of fatigue, there are a number of points worthy of note with regard to the 'simple performance decrement' approach to studying fatigue.

1.3.1.1 Implications for the conceptualisation of fatigue

The view of mental fatigue as analogous to physical fatigue implicitly conceptualises fatigue from a reductionist standpoint: While this is infrequently stated, this approach is most consistent with the assumption that repeated activity causes some impairment to, or deterioration of, a specific mechanism at a physiological level. There have been a number of explanations at this level, including that of Muscio (1921a) who stated that 'current scientific opinion' conceptualised fatigue as an accumulation of organic poisons in the active organ and that different degrees of fatigue would refer to different amounts of these activity products[2]. This 'toxin' model clearly corresponds to persisting models of physical fatigue, with regard to the build-up of lactic acid in the muscle. However, other explanations of mental fatigue at the physiological level also hold parallels with physical fatigue. For example, more modern, capacity based approaches to work and fatigue assume the existence of a hypothetical pool (or pools) of resources, comparable to reserves of physical energy, which are generally available for processing, but are essentially limited. Repeated mental activity is assumed to deplete the available 'resources' and it is the very lack of resources which causes the eventual performance breakdown.

There have also been a number of further explanations for physiological mechanisms which could underlie mental fatigue processes, including Tsaneva and Markov (1971) who argued that fatigue is determined by the production of a metabolite, which raises the synaptic threshold between a nerve and the organ that controls it, and Crawford (1961) who suggested that it is an increase in neural noise which underlies the effects of fatigue. This was regarded as a promising idea by Holding (1983), who argued that the changed signal-to-noise ratio could account for sensory fatigue effects, impeding decision processes, reaction times, fine motor control and heightened irritability and discomfort.

While these alternative models and approaches to studying fatigue obviously have different implications for the conceptualisation of fatigue as a unitary or multidimensional construct, there appears to be no discussion of such conceptual issues. However, in general, the repetitive application of a simple task is probably more consistent with a view

---

[2] It should be noted that Muscio made this statement with reference to existing models of fatigue, rather than offering support for such a conceptualisation.

of fatigue as a range of specific effects, rather than some general construct or central process: Consider again the analogy with physical fatigue – repeated stimulation of a specific muscle will quickly impair that muscle. However, it is assumed that there will be no 'general' effects, i.e. that the condition/performance of other muscles will not be impaired. On the other hand, if negative effects were transferred to other parts of the body which had not been stimulated, this would be more consistent with an idea of fatigue as a central process. Even with regard to 'mental fatigue', researchers have attempted to differentiate a range of different types. Ryan (1944) outlined seven different types of fatigue, for example the fatigue(s) produced by repetitive local tasks, emotional predicaments and problem solving. For each of the situations he suggested certain types of outcomes, supposedly signifying different kinds of fatigue. Furthermore, this trend continues today, as evident in Desmond and Hancock's (2001) differentiation of active and passive mental fatigue states.

While there may yet be some value this multiple fatigue approach, as these different fatigue states could be argued to differ on some level, it is possible to conceive of as many different 'types' of fatigue as there are fatigue generating conditions. This issue will be returned to in detail later. However, most relevant here is the fact that, while there may be some value in the reductionist conception of fatigue, which is implicit in the simple performance decrement approach – the methodology outlined in this section is without doubt wholly unsuited to addressing these theoretical questions.

1.3.1.2 Further limitations of the simple performance decrement approach

In addition to the problems of this approach outlined above, there are two further major limitations associated with adopting this approach to investigating fatigue.

Motivation as a moderator between performance and fatigue: The impact of varying levels of motivation has long been recognised in the fatigue literature. 'Motivation' was of course one of the central issues leading Muscio to his legendary conclusions – and, as an illustration of this, he included the following statement by Spaeth (1920; p204):

> *" if a man has been raising five or eight pounds for some time, he reaches a point where he can't work any longer, but if we removed a pound the individual finds he can go on. We also find, that when certain individuals are told, 'we are now removing a pound', but actually we add a pound instead of taking one off, the individual starts merrily on and proceeds to work"*

Another example of this effect was provided by the (much more frequently quoted) findings of Schwab (1953) who, in his investigation of muscle fatigue, found that subjects required to hang on a bar until 'exhaustion' would hang on longer if hypnotised, and longer still if offered a five dollar incentive to better their previous two tries! Motivation as a moderator in the fatigue process is obviously not confined to the effects of physical work, although these examples provide a very clear illustration of a psychological barrier to performance, rather than a physiological limit. Thus, the measurement of performance decrement as an indicator of fatigue is fundamentally flawed.

Furthermore, while the above examples provide an illustration of motivation as a limiting factor in performance, a high level of motivation can result in the maintenance of performance to an extraordinary degree. In both real life situations and experimental contexts, human beings have been recorded as maintaining performance under remarkable conditions and for astonishing lengths of time. In experimental settings, the best illustration of this phenomenon was provided by the classic experiments of Chiles (1955), who required participants to perform without rest for up to 56 hours. Even though some participants were physically unable to walk between testing stations, they continued to maintain their performance. This process, labelled 'performance protection' by Hockey (1993, 1997), will be discussed in detail in the next chapter. However, these illustrations further highlight the problems associated with employing simple performance measures as indicators of fatigue.

Limited view of performance: The final fundamental problem with this approach is the limited conception of performance. As previously stated, the analogy to physical fatigue had a substantial impact on the choice of research methods, resulting in the adoption of simple and repetitive tasks. Such tasks obviously possess inherently limited options for possible measures of performance, and the researcher is confined to assessing measures of only speed and accuracy. Adopting a broader view of performance has proved to be a much more successful approach to studying fatigue, but this has required substantial advances in the design of experimental tasks. (This will be discussed in detail in the following section). However, it is important to note that the emphasis on simple qualitative changes in performance has not only developed as a consequence of the physical fatigue analogy. Of further influence was the industrial perspective taken in a substantial body of the early fatigue research: As recognised by Cameron (1973) there was an extensive programme of research carried out by the Industrial Fatigue Research Board, during and after the First World War. This work was dominated by measures of productivity, and was even less likely to consider theoretical issues in studying fatigue, as the primary interest was the practical matter of output. Despite these different perspectives, the two strands of research were similar in that they emphasised impairments in performance as a consequence of repeated activity.

Therefore, the 'simple performance decrement' approach to studying fatigue is severely limited. Firstly, attempts to assess fatigue via performance impairment fail to account for the moderating effects of motivation and, secondly, its reliance upon simple tasks has limited options for measuring broader aspects of performance. Nevertheless, this approach has persisted. There remains a surprising volume of modern research which employs this approach and continues to discuss small impairments in performance as though they *are* fatigue – this is particularly true in fatigue research in sport, and will be returned to in Chapter 8.

1.3.2 Other performance based approaches

1.3.2.1 Performance on complex tasks

The second broad approach to studying fatigue is closely related to the methodologies described above, but takes a much broader view of performance. This approach again investigates fatigue from the perspective of the task-based consequences of carrying out prolonged mental work, but it is far more sophisticated in the selection of tasks, allowing for an entirely different level of performance analysis. The most celebrated and impressive example of such work was carried out by a team at Cambridge University during Second World War. Known as the 'Cambridge Cockpit Studies', this work was prompted by military aviation issues and utilised a complex flight simulator to investigate skill breakdown in pilots. A comprehensive description of the research group's findings is available in Bartlett's (1943) classic paper, based on his 1941 Royal Society Ferrier Lecture, but research findings are also presented in the independent reports of Drew (1940) and Davis (1948). In summary, the emphasis of this work was on the breakdown in the *organisation* of skill, rather than simple and gross measures of output. This breakdown in organisation had an impact on the internal timing of the successive elements and the disintegration of the task. This resulted in a shift away from a perception of a series of cohesive task components, towards the perception of a series of separate elements. Further features of this skill breakdown were the shifting standards in performance (operators unwittingly allowed progressively larger deviations of instrument readings before corrective action was taken) and a reduction in attention to peripheral items.

Interestingly, research considerations did not extend to the subjective feelings of those who took part (Cameron, 1973) although they did record a range of phenomena beyond the focus of task performance, including the reduced reliability of operator reports, the increasing awareness of physical discomfort and the tendency to blame errors on apparatus, rather than operators' own limitations.

This work was extremely enlightening, and represented a huge leap forward in our understanding of the complexities of human performance and the characteristics of performance breakdown. In comparison with earlier research methods, this complex-task approach was much closer to the requirements of real-world tasks, which are not typically of a simple repetitive kind, but are likely to involve an extensive hierarchy of sub actions relating to a series of (sometimes competing) goals.

Strategy shifts: In contrast to simple tasks, individuals typically have a whole variety of strategies which can be employed to reach the various goals which are characteristic of real-world tasks. As recognised by Bartlett and his colleagues, a shift in task strategy represents one of the most valuable on-task indicators of a fatigue state. In general, it is argued that fatigue will induce the individual to adopt changes in the way they carry out a task, employing methods which require a lower level of mental effort. In particular, recent research has focused specifically on the adoption of strategies which place less reliance on working memory (Hockey, 1986). An excellent example of this was provided by Sperandio (1978), who investigated changes in the strategies employed by air-traffic controllers. He found that, at certain levels of workload, controllers would switch from a bespoke plan to generalised procedures, reducing the load on working memory. This vein of research continues to be successful and will be discussed again in the next chapter.

Secondary tasks: A further aspect of performance breakdown that was evident in Bartlett's work was the narrowing of attention towards central aspects of the task. This was reported to occur in a physical sense, in that the stimuli which were peripheral within the display became increasingly ignored. This process is now widely recognised and formed the basis of Easterbrook's (1959) influential Cue Utilisation theory, which explains the effect of arousal on performance in terms of attentional narrowing. However, this process is also argued to occur in a goal-related sense, in that we are more likely to find performance impairments in aspects of a task which are lower down in the hierarchy of goals (Hockey, 1993).

Therefore, the complex task paradigm has proved extremely successful in advancing our understanding of skilled performance breakdown. Similar methodologies continue to form the basis of extensive research in the field of human factors, including a large body of work in the area of driving and the effects of performance breakdown on accidents.

Conceptual implications of this approach: Despite the undeniable value in this approach, as a method of investigating fatigue it is somewhat limited. The emphasis is on skill breakdown as a function of time on task. In Bartlett's classic paper (1943) entitled 'Fatigue following highly skilled work', Bartlett does not attempt to state what fatigue actually is. However, it is implicit within this work that fatigue *is* the disintegration or disorganisation of performance, as opposed to fatigue being a state which *results* in this pattern of performance breakdown. This approach highlights the concerns of Bartley and Chute (1947) who state that fatigue has been used more often as an explanation, than as something to be explained.

In Bartlett's later work, 'Psychological criteria of fatigue' (1953), he built on this perspective and defines fatigue as "a term used to cover all those determinable changes in the expression of an activity which can be traced to the continuing exercise of that activity under its normal operational conditions" (p1). It could be argued that this conceptualises fatigue at the purely specific level, referring to a breakdown of a specific process; possibly the process which is responsible for the organisation of skill.

A consequence of this approach, or view of mental fatigue, is that it provides no framework for considering a broader construct of fatigue, incorporating the conditions or states which follow sleep deprivation, emotional demands or physical work. Therefore, the performance decrements described by Bartlett and his colleagues may be better considered as a pattern of task deterioration associated with stress and high workload, rather than being an investigation of fatigue.

1.3.2.2 Probe tasks

An alternative performance-based approach to investigating the effects of fatigue is the use of 'probe tasks'. These alternative tasks can be either *interpolated* (inserted throughout the main or 'loading' task) or carried out at the end of the main task (most commonly known as 'after-effects tasks'). However, irrespective of their timing and frequency, they essentially adopt the same implicit conceptualisation of fatigue - that is, fatigue as a general or central state, rather than effects at the level of any number of specific processes. This approach is strongly supported by Holding (1983) and Hockey (1993, 1997) who argue that for the construct of fatigue to have any predictive value, it must be possible to show after-effects which extend beyond the loading task. Otherwise, it is difficult to argue that fatigue should be considered an entity, and could equally be conceptualised as nothing more than a collection of time-on-task effects, or collection of effects which result from physical work, emotional strain, etc.

This more theoretically driven approach also has a long history; for example, Muscio (1921a) used an interpolated task technique to investigate fatigue and found evidence of generalised effects of studying and attending lectures on subsequent speed of mental arithmetic. This finding provided evidence that there were some central effects of earlier mental exertions, which could influence a different class of mental operations. Furthermore, Fraser (1955) investigated the effects of long flights on a vigilance-based after-effects task. This task, carried out for one hour, required participants to recognise an occasional discrepant item in a display of otherwise identical items. While flights exceeding 10 hours reliably elicited a performance decrement, short trips (1-2 hours) were found to have no effect on task performance.

Other studies to find evidence of after-effects include the investigation of social behaviour by Cohen and Spacapan (1978) who found that helping behaviour decreased following a high level of real-life tasks carried out in a shopping centre, and Rotton, Olszewski, Charleton, and Soler (1978) who illustrated a lower level of persistence at impossible puzzles following high workload.

A detailed analysis of after-effects tasks, and the conditions under which they have been found to be sensitive, was provided by Broadbent (1979). However, he stressed that such after-effects have not always been easy to find. For example, Warren and Clark (1937) found little or no decrements on a range of mental and motor tasks, despite being kept on duty for up to 65 hours. In fact, the absence of performance after-effects could almost be considered a feature of fatigue research. However, rather than accepting this trend as evidence that fatigue is not a general state (but is confined to specific processes) Holding (1983) and others have put forward three generally accepted reasons for the characteristic unreliability of after-effects: Firstly, 'a change is as good as a rest' – it is argued that a direct consequence of changing task is that there is degree of change in participant state. Secondly, the evidence of any deficits will be masked, as motivated individuals are typically able to compensate, at least for short periods, and, thirdly, interpolated tests have simply attempted to measure the wrong kind of effect.

Aversion to further effort: With regard to the last of these points, there has been a growing trend to design after-effects tasks which allow for various levels of effort investment (Holding, 1983; Jongman, Meijman & de Jong, in press). Not only are such tasks consistent with the conceptualisation of fatigue as a central state, but their application emphasises 'aversion to further effort' as a central feature of this state. First identified by Thorndike (1900) and again highlighted by Bartley and Chute (1947), this approach contrasts with the implicit models of fatigue outlined above, in that it characterises fatigue more as a motivational state, rather than being some form of specific impairment, or breakdown in the organisation of skill. Also consistent with the view of fatigue as a motivational state is the recognition that the level of fatigue can be influenced not just by what one has already done, but also by what one has still to do! (Gaillard, 2001) Such a proposition is clearly at odds with the reductionist conceptualisation of fatigue.

Adopting this standpoint, Holding (1983) designed a series of after-effects tasks tailored to reveal the lack of willingness to expend further effort. These were based on Holding's strongly held belief that, other things being equal, fatigue subjects will choose to exert less effort. A further feature of these tasks is that 'lower effort' strategies are associated with a higher degree of risk. This is in accordance with the widely recognised phenomenon of an increase in risk-taking or carelessness when tired. Excellent illustrations of this have been provided by Brown, Tickner, and Simmons (1970), who identified an increase in risky overtaking by car drivers following a prolonged drive, and Webster, Richter and Kruglanski (1996), who found that participants were more likely to jump to conclusions and rely on stereotypes when tired.

In essence, Holding (1983) argues that tests of non-specific fatigue will be most sensitive if they present the individual with multiple strategies to achieve the goal, which are differentially weighted with regard to effort and risk (although it is not necessarily assumed that this process involves an explicit or rational evaluation of costs and benefits). This has been a particularly successful approach and has provided a methodological framework within which to investigate fatigue and address important theoretical questions, such as the relationships between the types of fatigue which result from different conditions and circumstances. In this vein, Barth, Holding and Stamford (1976) investigated the specificity of the after-effects following physical work, using a version of a task paradigm known as COPE (choice of probability/effort). In summary, they found that the effects of intense physical exertion had no after-effects on a perceptual version of the COPE task (which was based on the characteristics of a resistor-checking device). This suggests that the after-effects of physical work may not generalise to mental work, implying limitations to the generalisation of fatigue. However, in a motor-based version of the COPE task, 'physical' fatigue was found to be non-specific, e.g. leg fatigue influenced arm effort and resulted in risky choices when these promised to conserve physical effort.

1.3.3 Emphasis on subjective state

A markedly different approach to investigating fatigue is to consider it from the perspective of the individual. Subjective feelings of tiredness have long been recognised as central to the experience of fatigue, although attention to these feelings has been somewhat patchy: Some programmes of research have shown no regard whatsoever for subjective variables (e.g. the Cambridge Cockpit Studies). However, the recognition that evaluation of subjective state is a valid aspect of fatigue research is obviously evident throughout the literature. For example, Thorndike's work on the curve of 'satisfyingness' (Thorndike, 1917) and Poffenberger's (1928) work on the effects of 'continuous work on output and feelings' both afforded a central role to the subjective feelings associated with prolonged work. Furthermore, in 1946, a report from the US Civil Aeronautical Administration committee, headed by Viteles (Anon. 1946) categorically stated that performance effects should not be the exclusive area of interest in fatigue research and recommended a more human-centred approach. While the focus of the report was pilot fatigue, this report contrasted with the work of Bartlett and his colleagues, as there was a substantial shift in emphasis towards the individual. This approach to studying fatigue obviously has implications for the underlying conceptualisations. Unlike the physical fatigue analogy which is implicit within much of the simple task performance approach, or the emphasis on fatigue as a breakdown in skill organisation, the measurement of subjective state places the construct of fatigue at the subjective experience level.

The most influential work to emphasise the subjective perspective of fatigue was that of Bartley and Chute (1947). In their impressive volume "Fatigue and Impairment in Man" they differentiated between three facets of the fatigue problem; including the 'subjective feelings of tiredness and disinclination towards activity', 'impairment' and 'work decrements'. They defined *impairment* as specific changes at the tissue level, such as the accumulation of lactic acid in the muscles, and *work decrement* was defined as the overt activity that is measured either in the laboratory or in industry (e.g. task performance). However, while they accept that these two constructs are complexly related to fatigue, they argue that neither of these actually *are* fatigue. On the other hand, they argue that the subjective feelings of tiredness and aversion to effort should not merely be taken as part of the phenomenon, or as symptoms of fatigue, but as the fatigue itself. They state that, for the one who is tired, the feelings experienced are the fatigue.

As stated by Cameron (1973) this was a landmark monograph which provided a more thorough analysis of fatigue as an explanatory concept than anyone before or since! They hypothesised that changes in subjective state represent a biological warning that the individual's resources are overtaxed. However, this conceptualisation of fatigue as an adaptive 'stop emotion' did not originate here. In their rarely quoted publication, Whiting and English (1925) put forward a hypothesis of fatigue as a 'negative emotional appetite' which they differentiated from the physiological phenomenon of exhaustion: While exhaustion was primarily argued to be a physiological or specific and local phenomenon, they argued that fatigue is chiefly a subjective phenomenon. Like exhaustion it is a function of exertion, but unlike exhaustion, fatigue is not symptomatic of inability to work - they state that the relationship of fatigue to work is more complex. While this conceptualisation of fatigue had little direct impact on later research, it was acknowledged by Bartley and Chute themselves (1947) and is clearly consistent with their much celebrated analysis.

Implications of this approach: So, what are the implications of the conceptualisation of fatigue as a 'stop emotion'? Firstly, this is consistent with the probe task approach to studying fatigue, as fatigue is seen as part of the individual stance with reference to activity, whether that activity is vigorous physical activity, mental work or simply attempting to stay awake. An interesting question is to what extent these conditions or activities can be considered to generate the same single state. Bartley and Chute (1947) do in fact directly address this issue: they state that the feelings of tiredness and aversion are never specific to a given body member, never localised. They argue that, while bodily sensations can be localised, it is only the *individual* that can experience fatigue. These different bodily sensations associated with different activities are obviously different for the endless different origins of fatigue, however, they strongly state that the fatigue produced in one situation is essentially the same as the fatigue produced in any other: Having investigated different types of 'mental' and 'physical' fatigue, they argued that the differences are mostly those of localised bodily sensations, which differ from one activity to another. They provide the example of physical exertion, which they state leads to aches and pains of the gross musculature, as compared to mental exertion, which they argue leads to uncomfortable small muscles of the face, head and neck. However, this conceptualisation of a single central state of fatigue has been, by no means, universally accepted. Not only do researchers frequently adopt an implicit position that there are a number of different types of fatigue, but modern researchers have continued to state this explicitly. For example, Gaillard (2001) states that the fatigue which results from a demanding task that has been completed successfully is quite different from that resulting from a day of arguments and irritations. Of course, it would be difficult to argue that these conditions are different on some level, particularly with regard to the level of associated anxiety; however, whether the 'state of fatigue' is actually different depends on the nature and breadth of one's definition of fatigue.

The validity of these various hypotheses remains in question, as few researchers have attempted to address these fundamental theoretical issues. Nonetheless, the value of investigating fatigue (at least in part) from the subjective perspective cannot be underestimated. In common with the probe task methodology, this approach allows researchers to investigate the host of complex theoretical questions still surrounding the fatigue subject.

Chapter 2

A theoretical and methodological framework

2.1 Summary

This chapter outlines an alternative approach to the study of fatigue. In this theory driven approach, fatigue is explicitly conceptualised as part of a generalised response to stress. Central to this approach is the emphasis on a more holistic view of the individual, incorporating aspects of the wider context of performance, the biological response to stress and subjective experience. Adoption of this wider perspective is argued to allow a much more meaningful interpretation of fatigue research findings.

Accordingly, Hockey's (1993, 1997) Compensatory Control model of state regulation is adopted as a theoretical and methodological framework for the thesis. Some influences on the development of this model are outlined and the relevant aspects of the Compensatory Control model are discussed. This is followed by an examination of the implications of the model with regard to fatigue research methodology.

The end of this chapter provides an overview of the structure of the thesis and a description of the thesis parts.

2.2 Introduction

As discussed in Chapter One, fatigue research has a long history. Over the past hundred years there has been an ongoing practical and theoretical interest in the fatigue phenomenon. However, despite the vast body of research, there remains no generally accepted definition of fatigue and certainly no comprehensive model to explain the nature of the relationships among the different types of fatigue. The contrasting approaches to studying fatigue, outlined in Chapter One, have each emphasised different aspects of the problem and provided some insight into their respective variables of interest. For example, research into complex task performance has significantly advanced our understanding of the nature of performance breakdown following prolonged work and suboptimal conditions. This work has had a major impact in applied industrial contexts, in particular with regard to the design of safety critical systems. Another important development is the growing acceptance that *aversion to further effort* is a central element of the subjective experience of fatigue. This proposition also has many applications in the real world, including developing our understanding of risk taking behaviour and the tendency to shift towards risky decisions when fatigued. Therefore, fatigue research has proved to be both meaningful in predicting some patterns of behaviour and valuable in the guidance this information provides for the design of work and work environments.

However, fatigue research still lacks a comprehensive 'fatigue model'. While there is clearly some theoretical value in the methodological approaches outlined in Chapter One, this field suffers from the very circular problem that (a) there is no model of fatigue which could provide a framework to guide research and facilitate the interpretation of findings, and (b) these problems of interpretation further limit the development of an adequate model.

2.3 Fatigue as component of a generalised stress response

A significant step forward in developing a theory of fatigue was provided by Cameron (1973). In his paper entitled 'a theory of fatigue' he stated that the notorious difficulty of interpreting the results of fatigue studies was essentially due to the limited view of fatigue as a construct and lack of recognition that it is part of a broader process involving complex biological and motivational phenomena.

In common with Muscio, Cameron initially questioned the value of the fatigue construct and highlighted the wealth of well-documented problems associated with measuring fatigue. Furthermore, he challenged the recognised characteristics of experimentally 'fatigued' participants; such as performance breakdown, irritability and physical discomfort and argued these effects could equally be explained within the context of other phenomena, such as anxiety and sleep deprivation. However, Cameron then departed company with Muscio (and other researchers offering similar pessimistic conclusions, e.g. Browne, 1953) and stated that the fatigue problem cannot be made to disappear simply by defining it out of existence. His alternative approach was to view fatigue as part of a 'generalised response to stress' - by this he referred to the 'whole system of biological emergency mechanisms', and argued that the state of the individual can only be understood from the wider perspective of the individual's circumstances. He implicitly conceptualised fatigue as an outcome or product of the stress response by stating that the *degree* of fatigue experienced may depend on the level of the stress response as well as its duration. (Interesting, this implies that it is the duration of the stress response, not necessarily the duration of the stressful conditions, which is the important factor.) However, he maintained that fatigue is a construct with no explanatory value, as it is no more than a general description of a personal state at the time such changes are noted. (This idea will be returned to in Chapter 10.)

Despite the recognition of these issues, and the acceptance that his perspective fell short of a 'fatigue theory', he attempted to provide a theoretical groundwork from which to investigate fatigue. Central to this approach was the emphasis on a broader view of the individual, incorporating aspects of the wider context of performance, the biological response to stress, the subjective experience and individual differences.

2.4 A framework for fatigue research

Following the general groundwork provided by Cameron (1973), Hockey (1993, 1997) has further developed the proposition that fatigue should be considered as a component of the stress management response[1]. In common with Cameron, he argues that the adoption of this wider perspective of fatigue allows for a much more meaningful interpretation of research findings.

Accordingly, Hockey and Meijman (1998) propose that Hockey's Compensatory Control model of state regulation offers a suitable theoretical and methodological framework for future research into fatigue.

The following section will first outline some influences on the development of this approach and then summarise the relevant aspects of the Compensatory Control model as they relate to the fatigue construct (a full description of the model is provided elsewhere; see Hockey 1993, 1997). This will be followed by an examination of the implications of the model with regard to fatigue research methodology.

---

[1] This conceptualisation of fatigue and methodological approach precludes other types of physiological or pathological fatigue such as those associated with ME, low blood pressure etc. While there may be some commonalities, particularly on a subjective level, these types of fatigue are argued to be distinct and likely to arise through different mechanisms.

2.4.1 The Compensatory Control model

2.4.1.1 Background

The Compensatory Control (CC) model was developed as a cognitive-energetical framework for research on human performance. The emphasis is on the relationship between behaviour and its biological and motivational context and as such is argued to overcome the limitations of computational models of information processing, which fail to account for the variability of human performance under stress, emotion, and conditions that affect the general state of the individual. The incorporation of energy-based constructs (e.g. effort and resources) into the model allows some explanation of the different patterns of performance effects that have historically been observed under stress and high workload.

Similar attempts at integrating energetical constructs into information processing models have been made by Gopher and Sanders (1984), who attempted to integrate energetical mechanisms into a structural model of information processing stages, and Wickens (1986), who refers to the intensity aspects of information processing. However, Hockey argues that such approaches are somewhat limited in scope and provide little opportunity for investigating the complexity of performance across different categories of stressors, diverse environmental conditions and a broader range of tasks.

Therefore, this ambitious model aimed to be much broader in scope, with the wider remit of providing a comprehensive framework for research into a whole host of issues such as psychological health, strain, coping, individual differences, and fatigue.

Central to this model is the concept of compensatory mechanisms. This term was initially introduced by Freeman (1939) who was the first to use the term *compensatory behaviour* to account for the discrepancy between fatigue and performance in a work context. More influential though, was the work of Teichner (1968) who provided an impressive analysis of the interactions between behavioural and physiological stress reactions. In this review, he provided a very convincing argument for the importance of addressing compensatory mechanisms in stress research - as an illustration of compensatory phenomena he used the example of rectal temperature as a measure of the *stress response* to heat (p272):

> *"For example, the rectal temperature of men exposed to air temperatures of 100°F may show little increase. A naïve investigator observing only rectal temperature might be tempted to conclude that 100°F is not a stressor. In fact, at this temperature he could find an increased sweat rate, a raised temperature, an increased peripheral blood flow, and a decreased metabolic rate, all of which represent compensatory activities of the thermo-regulatory system."*

This example provides an excellent analogy to the use of simple performance measures as an indication of the underlying stress response and a very persuasive argument for the emphasis on compensatory processes in stress research.

2.4.1.2 Relevant aspects of the model

In common with other control models, a central assumption of the CC model is that behaviour is essentially goal driven. It is argued that individuals possess a dynamic hierarchy of both long-term and short-term goals. It is this internally-maintained goal state which directs action by setting the appropriate criteria for performance. Performance output is continually compared to the required standard and, where there is evidence of any mismatch between desired and actual state, these discrepancies can be 'corrected' so that performance is protected and stable over time. This maintenance of performance stability is considered to be an active process, i.e. under the *control* of the individual, which requires the management of cognitive resources through the regulation of mental effort.

Two levels of control: In common with the model put forward by Broadbent (1971), it is argued that this effort regulation occurs at two different levels of control. The lower level of control is responsible for routine corrections (small discrepancies between target and actual state). It is argued that this lower level of control possesses a set point or working effort budget, which is based upon the anticipated resource needs of the task. Increases in demand below this level are not felt as effortful and the control of performance therefore appears to be automatic. However, there is a second level of control (labelled supervisory or executive control) which is needed to deal with the regulation of effort when the discrepancy between the target and actual state is too great for low-level corrections to bring it within target range. When control of performance is shifted to this supervisory level, a number of optional modes of regulation are available, each with different consequences for task performance and associated costs. Broadly speaking, these conditions provide two options: (1) a downward adjustment of the goals, as a way of minimising the discrepancy between target and actual state, or (2) to respond by increasing the effort budget to accommodate the demands. The first of these options is categorised as a 'passive coping mode' (as opposed to the active coping mode of activity under lower level control). In passive coping, there is no adjustment of effort budget, and therefore this strategy would not incur compensatory costs, but obviously has implications with regard to the disruption of performance. As noted by Hockey (1997) this passive mode of control corresponds closely to Frankenhaeuser's (1986) 'distress without effort' pattern of coping. Frankenhaeuser focused on neuroendocrine activity as an aspect of the general stress response, and outlined patterns of activity associated with different coping modes. Her 'distress without effort' pattern is characterised by a marked increase in cortisol production, and a modest increase in adrenalin. This is argued to represent feelings of helplessness and the loss of control.

The alternative to passive coping within the current framework is the 'strain coping mode', which is characterised by an increase in compensatory effort. However, there are expected consequences of such regulatory activity, in the form of emotional and physiological costs. In Frankenhaeuser's framework this corresponds to the 'effort with distress' pattern of coping, which is characterised by a significant increase in adrenalin and, to a lesser extent, cortisol (in addition to the subjective experience of anxiety and fatigue).

Compensatory control and fatigue: Therefore, within the framework of this model, fatigue is hypothesised to occur as a direct consequence of the adoption and maintenance of supervisory level control - it is argued that there is a direct relationship between the application of compensatory effort and the development of fatigue.

2.4.1.3 Implications for research methodology: patterns of latent degradation

In summary, the Compensatory Control model argues that individuals may defy normal expectations of performance because of the presence of compensatory control mechanisms, which protect primary task activities. This has significant implications for the study of fatigue as any investigation must look beyond simple changes in performance, and consider a broader range of parameters. These parameters are identified following the predictions of trade-offs between effort and costs, which represent the very foundations of this theoretical model.

As fatigue is hypothesised to occur only when an individual engages in the supervisory-level control of goal-driven behaviour, the effects of fatigue should only be detectable in secondary tasks, after-effects and subjective state, rather than performance breakdown on the primary task. Within the context of goal-directed behaviour, primary task performance is unlikely to deteriorate unless the conditions are sufficiently extreme. Thus, in simple terms, an individual working in a high workload environment has the opportunity to accept or reject their current goal. If they accept the goal, a *strain* pattern of decrement is likely, which would involve maintenance of the primary task, at the expense of secondary tasks, performance after-effects and associated changes in subjective state (high mental effort and increased ratings of fatigue). Alternatively, while performance may be protected on the primary task, an individual may adjust their primary task management strategy and adopt simpler strategies.

The other option is for individuals to reject the primary goal, which would result in a *passive* coping strategy. In such circumstances, a different pattern of performance is predicted, in which performance decrements may be observed on the primary task. However, this strategy is unlikely to generate a fatigue state within the individual, as it is characterised by low effort, therefore, there would be little impact on subjective measures of fatigue and after-effects tasks.

These predictions are extremely important in directing fatigue research and in facilitating the interpretation of research findings. The observation of these different patterns of performance breakdown led Hockey (1998) to argue that any investigation of the fatigue phenomenon should go beyond the consideration of primary task performance (which is unlikely to be a consequence of fatigue) and instead focus on patterns of latent degradation, which essentially provide measures of convergent validity. For example, if there is a level of deterioration on primary task performance, despite the adoption of a high effort strategy, it should be possible to confirm this with evidence from sources such as subjective assessments or psychophysiological measures.

The four categories of latent degradation which are identified within this model have been well documented (Hockey, 1993; 1997; Hockey & Meijman, 1998 etc) and will be discussed again in Chapter Five. However, in brief, they are: *secondary task decrements* (involving selective impairment of lower priority task components), *changes in task management strategies* (involving within-task shifts to simpler strategies), *regulatory costs* (defined as the unwanted side-effects of regulatory activity - e.g. increased mental effort and sympathetic dominance) and *after-effects* (post-task preferences for lower effort strategies and subjective fatigue).

Interestingly, the possible forms of latent degradation suggested by Hockey correspond very closely with the individual approaches of studying fatigue outlined in Chapter One. The previously discussed methodological approaches of complex task performance, probe tasks and subjective state assessment are all incorporated into the current theoretical framework. However, Hockey argues for a holistic approach which takes into account the biological/motivational context of goal-directed performance and makes clear predictions about patterns of performance breakdown following the adoption of various control modes. Without this broader view, experimental manipulations of fatigue would be extremely limited and assumptions about the presence or absence of fatigue may be invalid. In such instances it would be to draw meaningful conclusions about the nature of the fatigue state, and address important theoretical questions.

2.5 Plan of the thesis

This thesis investigates a range of theoretical issues surrounding the fatigue construct, in particular the nature of the relationships between the different types of fatigue.

The work is divided into four parts and contains ten chapters in total:

Part One contains two chapters which provide an introduction to the central issues of the thesis and the relevant background research: Chapter One specifically addresses the different approaches which have been employed to study fatigue and the conceptual implications of these approaches. Chapter Two outlines a theoretical model which is used to guide the experimental work and facilitate the interpretation of the findings.

Part Two comprises Chapters Three and Four and focuses on the subjective and psychometric basis of the fatigue construct. While the theoretical model outlined in Chapter Two provides an excellent framework within which to carry out experimental fatigue research, this model does not make any predictions with regard to the dimensionality of the fatigue construct. Therefore, before undertaking the experimental work, an investigation of the subjective and psychometric bases of the fatigue construct was required. Chapters Three and Four utilise a range of methodologies to examine individual experiences of fatigue, and the subjective and psychometric evidence in support of a unitary or multidimensional construct. This work provides some insight into the relationships between the different types of fatigue which are evident in the existing literature. This leads to the development of two multidimensional research scales of state and trait fatigue. The state fatigue scale is later utilised to assess the pattern of changes in fatigue following different kinds of work.

Part Three presents the experimental components of the research, based on the theoretical and methodological framework outlined in Chapter Two. Central to this work is the hypothesis that fatigue can be conceptualised as a component of the stress management response and that it develops in response to effortful goal-directed activity. Therefore, different types and different levels of work are experimentally manipulated to facilitate an investigation of the dynamic development of fatigue. Chapter Five outlines the experimental methodology, which describes the important features in the design of the following experiments - in particular the computer based simulator which provides opportunities to examine patterns of performance breakdown and evidence of compensatory activity. It is from this multilevel perspective that the development of fatigue is investigated. While Chapters Six and Seven focus on the development of fatigue following mental work, Chapter Eight considers the development of fatigue following physical work. Chapter Nine represents a culmination of all the work presented thus far and considers the nature of the interaction between mental and physical fatigue.

Part Four contains only one chapter and provides a summary and general discussion of the findings and issues considered throughout the thesis. This chapter concludes with a discussion of the conceptual implications of the findings for a comprehensive fatigue model.

Part 2:

The Subjective and Psychometric Basis of the Construct of Fatigue

Chapter 3:

A Preliminary Investigation of the Fatigue Construct

3.1 Summary

The central question running throughout this thesis is whether fatigue should be considered as a unitary or a multidimensional construct. Although a number of different types of fatigue are recognised, it is unclear whether these represent truly different experiences. Furthermore, the literature is plagued with inconsistencies regarding the 'types' of fatigue which should be incorporated into a conceptual framework. Therefore, rather than basing the research programme entirely on existing research, the work embarked upon in this chapter adopted a more grounded approach (Glaser, 1967).

This chapter describes the preliminary research which uses interviews and questionnaires to investigate personal experiences of fatigue. Ten in-depth interviews were carried out which identified a broad range of circumstances believed to generate fatigue and different subjective experiences of the fatigue condition. This work provided some limited support for the hypothesis that it is possible to differentiate between the subjective experiences of different 'types' of fatigue, as well as identifying two distinct strands of possible enquiry. These relate to state fatigue and trait fatigue. The remainder of the chapter then focused on trait fatigue and, building on the findings from the interviews, a psychometric investigation of the fatigue construct was carried out.

Ninety-eight fatigue-related items were generated, which incorporated the full range of conditions identified as potentially giving rise to a fatigue state. Exploratory factor analysis revealed six distinct factors of susceptibility to the development of fatigue under different conditions. This provided strong support for the proposition of a multidimensional construct of trait fatigue.

3.2 Introduction

As discussed in Chapter One, fatigue research has not enjoyed the advantages of an established theoretical framework on which to build, and thus our understanding of the fatigue construct remains limited. Consequently, the fundamental question of whether fatigue should be considered as a unitary or a multidimensional construct continues to be unresolved: Is there a single general process of fatigue, generated by a range of different environmental conditions, or a number of distinct and separate mechanisms? This question clearly raises important practical and theoretical issues, particularly in terms of the joint effects of the different possible sources of fatigue: Modern working environments typically impose different demands which may generate more than one type of fatigue.

Implicit in the research literature is the general assumption that there is more than one type of fatigue (physical and psychological fatigue being the most frequently and easily differentiated). Although few would disagree that these two 'processes' differ on some level, it is not untypical to go beyond this and differentiate between other types of psychological fatigue on the basis of their different sources, e.g. sleep-related fatigue as opposed to the condition which develops in response to cognitive demands. Further complicating this issue is the lack of agreement on how many different 'types' of fatigue there are. Even if we assume that there are a number of different fatigue processes, there is still no consensus on which types of fatigue should be incorporated into a conceptual framework. As stated in Chapter One, the types of fatigue considered have typically depended on specific research aims (e.g. research into the effects of shiftwork may consider 'sleep-related fatigue'; research into fatigue in a caring profession may focus on 'emotional fatigue'). But, this lack of systematic enquiry has done little to develop our understanding of the relationships between the different fatigue types.

These issues are of primary importance in the development of a fatigue model. Therefore, the work discussed in this chapter represents a preliminary attempt to investigate the range of possible types of fatigue which should be considered. While the theoretical model discussed in Chapter Two provides an excellent framework within which to investigate fatigue experimentally, the model offers no predictions regarding the dimensionality of the fatigue construct. Rather, it offers an approach suitable to addressing such questions. However, before the experimental work was undertaken, a subjective and psychometric investigation was carried out. This work would then inform the experimental work, and provide appropriate subjective assessment tools.

3.3 Preliminary Field Research

Existing fatigue research certainly provides a rich source of information on which to base an investigation into the nature and 'breadth' of the construct. But, although there is a substantial body of literature which considers many different types of fatigue, this literature is plagued with inconsistencies. Hence, past research and existing questionnaires may not represent a comprehensive base on which to consider the range of experiences which could be incorporated within a theoretical framework. On this basis a grounded approach was adopted here, which borrows from the perspective of Glaser (1967). This approach emphasises the emergence of theory from unstructured data collection, rather than the restriction of data collection on the basis of pre-existing theory.

A small exploratory study was carried out which aimed to investigate a broad range of personal experiences of fatigue and lexical distinctions for different types of feelings. This methodology borrows from early attempts to investigate the structure of personality (Cattell, 1947): In everyday language we make distinctions between different types of tiredness e.g. *exhausted* from a hard day at work, *drained* as a result of dealing with emotional problems or *sleepy* as a result of sleep deprivation. These different terms, used to talk about tiredness, provide a rich source of data which may represent an intuitively sensible starting point for fatigue research. Not only would this approach facilitate a grounded exploration of the range of different types of fatigue, but one could also investigate the existence of any subjective distinctions between these fatigue types.

3.3.1 Interviews

A series of ten in-depth semi-structured interviews were carried out with three psychologists and seven undergraduate psychology students. The aims of this process were (a) to identify individual notions of fatigue (b) to identify circumstances and events which generate fatigue (c) to investigate different experiences of the fatigue condition. To meet these aims, a flexible interview schedule was designed which incorporated a range of components:

1. Individual notions of fatigue

To identify individual notions of fatigue, interviewees were asked general open questions such as "What do you think of when I say the word fatigue?" and "What does this word mean to you?" This questioning generated a fairly narrow range of fatigue synonyms e.g. tired, exhausted, lethargic, drowsy, drained etc. All ten interviewees began by stating that fatigue was "tiredness" or "being tired". This supports existing definitions of fatigue, which typically agree that the construct is characterised by a subjective state of tiredness. In addition to the use of twelve fatigue synonyms (see Table 3.1) one participant defined fatigue by its opposite, stating that fatigue is "when you don't have much energy" and two interviewees referred to fatigue in terms of the condition which follows on from work, stating that it is how you feel after "a hard day's work". Therefore, the subjective notions and personal models of fatigue were fairly limited.

Fatigue Synonyms			
Tired/tiredness	(10)	Drowsy	(2)
Exhausted	(9)	Whacked	(2)
Shattered	(6)	Lethargic	(1)
Knackered	(6)	Done in	(1)
Sleepy	(4)	Flagging	(1)
Drained	(3)	Sluggish	(1)

Table 3.1: Fatigue synonyms derived from preliminary interviews (Numbers in brackets refer to frequency of interviewees using this word)

2. *Generators of fatigue*

Following on from the above questions, interviewees were asked to identify the different conditions which might lead to or generate a state of fatigue. This questioning produced a long list of potential antecedents, the most common of which were various types of mental work. All participants referred to tasks requiring 'concentration', 'prolonged thinking' or some form of mentally demanding task. Surprisingly, mental work was more frequently quoted than examples of physical work, such as sport and exercise, physical labour, or shopping. In addition to mental and physical work, emotional difficulties, boredom, poor quality sleep, travel, illness and the extremes of weather were all quoted as conditions which could lead to fatigue. Interestingly, there appeared to be some substantial individual differences within the small sample interviewed, not only in terms of the specific conditions which generate a state of fatigue, but also in the regularity with which people feel tired. Some participants stated that they regularly experienced fatigue, while for others this was quoted as a fairly infrequent state. This supports the hypothesis (discussed further in section 3.4) that fatigue can be considered as a trait, as well as a state resulting from certain conditions.

3. *Experiences of fatigue*

Having identified a broad range of conditions which could potentially induce fatigue, Critical Incident Technique (Flanagan, 1954) was used to explore specific personal experiences in greater depth. Interviewees were asked to recall a time in which they had been fatigued. This 'incident' was then analysed in terms of the events leading up to the experience, the nature of the ensuing state, and the recovery. Each interviewee selected a minimum of two incidents and a maximum of four (dependent on the depth with which they could discuss each incident). These included physical work, sleep problems, and emotional conflicts. However, most frequently, incidents were based on intense mental demands, either at work or during student exams. Interestingly, all selected fatigue-related incidents involved some component of anxiety, with the exception of the state following physical work. The physical incidents were also distinct from other fatigue generators, because of the inclusion of physical symptoms, such as muscle aches. (Although the presence of physical symptoms in mentally fatigued individuals has long been recognised; e.g. Bartlett, 1943). Despite these differences, the terms used to describe the experience of fatigue were quite similar. Terms such as 'tired', 'exhausted' and 'knackered' were used equally to describe the type of fatigue which results from either physical work, mental work or emotional work. These terms seemed to capture the essence of the subjective experience of fatigue, irrespective of the conditions which generate the state. However, for some incidents, it was possible to identify patterns in the way certain terms were used. The most obvious of these were the terms used to describe incidents related to sleep disturbances, such as night work and new parenthood: When describing such incidents, interviewees tended to use words such as 'sleepy' and 'drowsy'. More interesting were the terms used by two interviewees to describe their feelings following emotional problems. Although the terms they used to describe their state were similar to those used to describe other incidents, they tended to be prefaced by the word 'emotionally', e.g. 'emotionally tired' and 'emotionally drained'. When questioned further, both interviewees for whom this had been the case said it was possible to differentiate between the feeling of being emotionally drained and physically or mentally drained, although only one could qualify this further, stating that emotional fatigue was associated with the feeling of emptiness. This provides some support for the hypothesis that there is more than one type of fatigue, as the subjective experience appears to differ on some level. However, it also emphasises the impenetrable nature of the subjective experience of fatigue.

Although the interviews raised many questions, they also met their broad goal. This was to act as a preliminary investigation into the subjective experience of fatigue, and to evaluate the breadth of conditions which can give rise to fatigue-like states. Some of the conditions identified by the interviews went beyond those previously identified in the literature, such as the feelings of fatigue associated with travel, extreme temperatures and alcohol. Although some of these fatigue generators could be described as a little obscure, they may prove significant in the broader investigation of the fatigue construct, and it was considered important to begin with a comprehensive and grounded investigation. Furthermore, the interviews confirmed that is it is possible to differentiate between the subjective experiences of the fatigue which is generated from different sources.

3.3.2 Further investigation of fatigue synonyms

Although the interviews provided some evidence that it is possible to differentiate between the subjective experiences of different types of fatigue, this approach met with limited success. As previously mentioned, fatigue appears to be an extremely commonplace, yet impenetrable, experience. However, if the distinction between the different kinds of fatigue is more than simply operational (produced by different conditions), this should be measurable using a questionnaire method and factor analysis to determine the underlying factor structure (Hockey, 1996).

3.3.2.1 Dictionary search

The preliminary interviews provided a basis for a psychometric investigation of the construct: In addition to the broad range of fatigue generators identified by the interviews, an array of terms which could be used to describe the state of fatigue was also identified. However, this was based on a small sample. Therefore, to ensure that the 'fatigue domain' had been adequately sampled, the Oxford English Dictionary & Thesaurus were systematically searched. The terms generated by the interviews were used as a starting point to provide fatigue synonyms, and each of these cognate words then provided the basis for the next level of search. This search, which continued until no new words were found, generated ninety-seven fatigue-related words.

3.3.2.2 Word usability analysis

Not surprisingly, a number of the fatigue-related words generated from the dictionary search were very unusual, and may be unknown to a substantial proportion of the general population. Although a thorough investigation of the fatigue domain is essential for content validity, it is also important that any psychometric investigation is based on familiar words, as the inclusion of unfamiliar words may influence the resulting factor structure. Before further investigation could proceed, it was considered important to carry out some analysis on the fatigue-related words, as it would not be appropriate to incorporate unfamiliar terms. To assess the usability of the words, a small survey was carried out: A document containing the ninety-seven fatigue-related words was distributed to fifteen individuals (five academic psychologists, five students and five office workers). Respondents were instructed to (1) *strike through any words with which they were unfamiliar* and then (2) *highlight any of the remaining words which they did not feel were related to fatigue or tiredness.* This approach follows that adopted by Spielberger (1983) who reduced a pool of 177 anxiety-related items to 124, eliminating those which were rated as redundant, vague, or ambiguous.

From the original 97 words, twenty-nine were removed on the basis of unfamiliarity. These included 'enervated', 'supine', 'fallow', 'lassitude' and 'torpid'. There was a surprising level of consensus between raters and, with the exception of one word, all words that were rated as unfamiliar were excluded from further analysis ('listless' was identified by one rater, but was considered to be a sufficiently familiar word to be retained). Of the remaining sixty-eight words, a further twenty-three were identified as being ambiguous or unrelated to fatigue. These included 'impotent', 'smashed' and 'vacant'. A total of forty-five words remained.

3.4 Psychometric Evaluation

The preliminary field research provided a rich source of data upon which to base further investigation of the fatigue construct. The interview process and dictionary search provided two distinct (although closely related) types of data: Firstly, there was data regarding the circumstances or conditions in which individuals report having developed fatigue and, secondly, lexical descriptions of the subjective experience of the fatigue condition.

These two types of data are related to two distinct channels of fatigue investigation, namely short-term (state) fatigue and longer-term (trait) fatigue: Trait fatigue refers to the relatively stable individual differences in the predisposition to experience fatigue, whereas state fatigue is defined as a transitory reaction or process taking place at any given time and level of intensity.

A thorough investigation of both state and trait fatigue is essential in developing our understanding of the fatigue construct. However, while later work focuses on state fatigue and the development of fatigue under specific conditions, the remainder of Part 2 focuses on trait fatigue as a psychometric investigation of this aspect of the fatigue construct provides an excellent starting point to address the question of whether fatigue should be considered as a unitary or a multidimensional construct: If there is a single general process of fatigue, then individual predisposition to fatigue should be consistent across a range of environmental conditions, resulting in a single fatigue factor. However, if fatigue is more correctly conceptualised as a number of distinct and separate mechanisms, then individuals may be differentially predisposed to fatigue following different environmental conditions. If this is the case, factor analysis would reveal a number of fatigue factors. Some evidence towards this question is already available in the form of existing fatigue scales and questionnaires. However, there are some fundamental limitations to this research, which makes firm conclusions problematic.

3.4.1 Existing Questionnaires

There are a substantial number of fatigue scales in existence. However, the factors or fatigue components included in them varies widely between scales, as they have typically been developed to measure the specific 'types' of fatigue which are relevant to the researchers. For example the Chalder Fatigue Scale (Chalder, Berelowitz, Pawlikowska, Watts, Wessley, Wright *et al*, 1993) revealed two distinct factors of physical and mental fatigue when subjected to factor analysis. However, the questionnaire items were initially generated by various experts with the *specific* aim of reflecting physical and mental fatigue. This circular process of writing items for assumed factors and then producing a clear factor solution has been echoed in the development of many other scales, which will be discussed further in Chapter 4.

A number of other researchers have identified distinct fatigue factors (these will be discussed in detail in Section 4.3). However, in common with Chalder and her colleagues, their scale items have typically been written with specific types of fatigue in mind; either on the basis of research requirements, or on assumptions about the nature of the fatigue construct. No scales have considered the full breadth of possible types of fatigue, a process which is necessary for a comprehensive investigation of the construct.

3.4.2 Preliminary Scale Development

To fully investigate the underlying factor structure of trait fatigue (or individual predisposition to fatigue in various conditions) it was considered necessary to develop a new questionnaire. A broad range of items were generated which consisted of a statement followed by 6 point Likert scale with verbal anchors ranging from (1) *strongly disagree* to (6) *strongly agree.* The item content was based on the data provided by the preliminary field work. The full range of fatigue generators was utilised, comprising items relating to mental work, emotional 'work', physical work, boredom, sleep patterns and disturbances, weather, travel, health and food. Also, a range of appropriate fatigue synonyms were incorporated into the statements to reduce the repetition of the words *tired* and *fatigue,* for example "Listening to other people's problems leaves me feeling *drained*" and "I feel *worn out* after exerting myself physically". The selection of these synonyms was limited to those which were generated by the Critical Incident interviews to minimise the confounding of states and verbal descriptions. The initial questionnaire consisted of 98 items which all complied with the guidelines for item writing set out by Kline (1986). This ensures items are clear, unambiguous and specific.

The items were then piloted with ten respondents who checked each item for clarity and ambiguity. All items were considered to be clear and of suitable language for the general population. However, following feedback from the pilot, a change was made to the instructions regarding the apparent repetition of similar items: two of the respondents commented on the similarity of some items and so it was considered important to make the following statement in the instructions. “Although some of the statements are similar to others please respond to each one separately”. It was considered that this warning would reduce the tendency to look back and respond to similar items in the same way. (A copy of this questionnaire is available in Appendix 1.)

3.4.3 Sample Characteristics

Four hundred and fifty-two questionnaires were completed and returned from seven hundred distributed. This represented a response rate of 64%. Due to easy availability of students, this group constituted 65% of the sample (n=295). However, other occupational groups were actively sought out. Of the remaining 157 respondents, 71 (15.7%) were nurses, 42 (9.3%) were physiotherapists, 26 (5.7%) were secretarial and 18 (3.9%) were other professionals. The sample consisted of 292 (64.6%) females and

139 (30.8%) males. (21 respondents did not state their gender.) The age range of the sample was 18 to 65 (mean = 25.64; sd = 9.65).

3.4.4 Data Treatment and Results

The data were coded, entered into SPSS and subjected to data reduction. Principal components analysis with varimax rotation was used and eigenvalues greater than one were accepted. A scree plot was examined to help determine the possible number of factors before rotated solutions were studied.

In the initial solution, 27 factors had eigenvalues greater than one and accounted for 68.6% of the variance. Not surprisingly, this solution possessed little psychological meaning. Examination of the scree plot suggested that a six or seven factor solution may offer a superior description of the data. However, to fully investigate the latent factor structure, four solutions were explored ranging from five to eight factors. Varimax rotation was used to aid interpretation. The five factor solution accounted for 32% of the total variance. The first rotated factor, accounting for 9.4% of the variance, did not produce a very coherent cluster. This included items relating to fatigue in response to prolonged mental work, emotional conflicts and some of the more obscure items such as sensitivity to extremes of weather and long journeys. Although the remaining four factors represented slightly more coherent clusters relating to physical fatigue and three sleep related factors, these factors also contained other items. For example the physical fatigue factor (accounting for 6.9% of the variance) also contained items relating to general energy/vitality. Therefore, the five factor solution did not seem to provide a meaningful description of the fatigue construct.

The six factor solution accounted for slightly more of the total variance (35.1%) and represented a more satisfactory solution (see Appendix 2). The first rotated factor, accounting for 8.2% of the variance, was again the least coherent, incorporating items from almost all domains generated by the field work. This included tiredness as a result of dealing with emotional conflicts, mental work, physical work, driving and extremes of weather. However, on close examination of these items, this cluster could be argued to support the general notion of **fatiguability** as identified by Flugel (1928) who noted "marked individual differences in fatiguability, just as there are in ability, in improvability and in daily variability". That is, that some people are simply more prone to developing fatigue as they go about their day-to-day activities than others. The second rotated factor, accounting for 6.8% of the variance, comprised all of the **physical fatigue** items, such as *"I feel worn out after exerting myself physically"* and *"being physically active makes me mentally alert"* (negative loaded). These items had relatively high loadings on factor two, ranging from .529 to .710, suggesting a strong physical factor. Interestingly, factor two also had a subcomponent of '**energy**' items, such as *"I expect to feel lively"* and *"I come across to others as being a bit lethargic"* (negative loaded). Furthermore, these items did not load on to factor 1, suggesting a distinction between general feelings of energy and vitality (factor 2), and the predisposition to develop fatigue following various activities (factor 1). Factor 3 also offered a coherent and meaningful cluster. All but two of the eleven items clearly related to **mental fatigue**, such as *"I get worn out by prolonged mental activity"* and *"I get tired when I need to concentrate hard for long periods"*. Although there were two distinct items which were not instantly recognisable as mental fatigue, they were theoretically consistent with this cluster, as they both referred to remaining alert throughout the day and possessed negative loadings. Therefore, these could be considered as negative mental fatigue items, as they relate to the absence of a fatigue build-up when dealing with daily work. The remaining three factors were all sleep-related. Factors 4 and 5 related to **morning tiredness** and **evening tiredness**, and accounting for 5.2% and 5.0% of the variance respectively. Interestingly, these two factors closely resemble the 'owl and lark' traits identified by Horne and Ostberg (1976): Factor 4, morning tiredness, contained items such as *"I take a long time to get going in the morning"*, *"I prefer to do difficult work in the afternoons"* and *"I am full of energy first thing in the morning"* (negatively loaded). This is closely related to Horne and Ostberg's owl, whereas factor 5, evening tiredness, is closely related to the lark. This factor contained items such as *"I am full of energy first thing in the morning"*, *"I wake up feeling tired"* (negatively loaded) and *"I have trouble getting to sleep at night"* (negatively loaded). Factor 6 also related to sleepiness, as all of the items referred to **nap taking**. This was again a very coherent cluster of items, which accounted for 4.1% of the variance.

Therefore, the six factor solution offered a meaningful and interesting description of the data. However, the seven and eight factors solutions were also investigated to ensure that six factors offered the best reduction of the data. The seven factor solution, accounting for 37.0% of the total variance, provided a less clear solution than six factors. Four of the factors provided little coherence, containing items related to mental fatigue, emotional fatigue, sleep-related fatigue and, in some cases, physical fatigue. Only factors 2, 4 and 5 offered meaningful clusters. Factor 2, accounting for 7.1% of the variance, echoed the physical factor from the previous solution. This contained items relating to fatigue following physical activity and also energy-related items as above.

Factor 4, accounting for 4.9% of the variance, contained items similar to the fatiguability factor identified by the six factor solution and factor 5, accounting for 4.3% of the variance, contained only items relating to nap taking.

The eight factor solution, which accounted for 39.2% of the total variance, provided a similar picture to the seven factor solution. There were some clear and meaningful clusters, while others provided little coherence. The clear factors related to physical fatigue, nap taking and morning and evening tiredness and there was also a fairly coherent mental fatigue factor. However this 'mental' cluster also contained some items relating to emotional fatigue and general alertness. The remaining three factors were very similar and contained a range of items from almost all of the possible domains.

Therefore, the six fatigue factors appears to offer the most meaningful and coherent solution. However, of possibly greater significance is the robust nature of some of the clusters identified, a number of them being present in the five, six, seven and eight factor solutions. In each solution, there was little overlap between mental and physical fatigue items and the sleep-related items. The least robust of these was the mental fatigue factor, which clustered with emotional items and other more general fatigue generators in the five factor solution and also in the seven and eight factor solution.

Only the six factor solution provided a very coherent cluster of mental fatigue items. Interestingly, the mental and physical fatigue items consistently loaded on to different factors in each solution investigated, which supports the notion that these fatigue processes differ on some level. Furthermore, this psychometric investigation also supports the theorists who go beyond two types of fatigue and differentiate between other types of psychological fatigue, e.g. sleep related fatigue as opposed to mental fatigue.

In conclusion, this subjective and psychometric investigation of the fatigue construct provides strong support for a multidimensional model of fatigue. Not only is it possible to differentiate between the subjective experiences of the types of fatigue generated by different conditions, but also the predisposition to develop fatigue across various conditions appears to differ between individuals.

Chapter 4:

State and Trait Fatigue Questionnaires

4.1 Summary

Chapter 4 again adopts a psychometric approach to addressing the conceptual issue of dimensionality of the fatigue construct. Evidence in support of a multidimensional construct was found in existing fatigue questionnaires and in the psychometric investigation of both trait fatigue and state fatigue. The factor analysis of the trait fatigue questionnaire revealed clear and seemingly robust distinctions between physical, mental and general fatigue and a range of sleep-related factors. This provided initial support for the notion of a number of distinct underlying mechanisms which differentially predispose individuals to develop fatigue in specific given circumstances. This factor structure was further supported by construct and criterion related validity evidence.

As tentatively predicted, the factor analysis of the state fatigue data did produce a different factor structure from that of trait fatigue. Distinctions were again found between state mental, physical and sleep-related fatigue. However, the factor analysis of the state data did not produce the same clarity of distinction between these different types: The initial factor analysis produced some level of distinction between the fatigue scales and negative affect and boredom scales, but limited evidence for differentiating between physical, mental and sleep-related fatigue. Although the validation process of the state fatigue scale provided good support for the differentiation between the types of state

fatigue, the distinction between mental and sleep-related fatigue was not clear cut and, in general, the support for a multidimensional concept of *state* fatigue was less convincing than for the concept of *trait* fatigue.

4.2 Introduction

Chapter 3 outlined preliminary research into the subjective and psychometric basis of the fatigue construct. In line with existing research, this analysis provided support for a multidimensional construct of fatigue. However, this construct clearly requires further investigation. This chapter continues the exploration of the psychometric basis of fatigue on the two distinct levels of state and trait fatigue, and outlines the evidence in support of a multidimensional construct. However, before describing the psychometric investigation, this chapter will first provide a detailed evaluation of existing fatigue scales, identifying important feature of the various scales and outlining the broader implications of the disparate approaches.

4.3 Existing Fatigue Questionnaires

Although a broad range of fatigue questionnaires have been developed, few have addressed the fundamental issues surrounding the nature of the fatigue construct. As stated in Chapter 3, attempts to measure fatigue have typically been born of practical necessity and research requirements, rather than theoretical enquiry. This has led to a long series of bespoke instruments with different aims, different (or no) theoretical underpinnings and, not surprisingly, different findings. But, despite this largely piecemeal development, many of the bespoke scales have been well developed for their particular requirements, thus offering some valuable evidence for theoretical developments.

4.3.1 Historical perspectives

The complexity of fatigue measurement has long been recognised. In 1928, Poffenberger succinctly summarised what remain the key issues in measuring subjective fatigue in his paper "The effects of continuous work upon output and feelings". Although there had been earlier attempts to measure the feelings associated with mental work, such as Thorndike's work on the curve of 'satisfyingness' (Thorndike, 1917), Poffenberger's paper represented the first real attempt to address the specific issues associated with fatigue *scales*. As he was interested in measuring tiredness associated with mental work, he devised four different mental tasks to generate fatigue experimentally. These included continuous addition, sentence completion, judging compositions and intelligence tests; each continuing for approximately 5½ hours. Feelings were recorded on a one-dimensional seven-point verbally anchored rating scale, ranging from feeling (1) 'extremely good' to (7) 'extremely tired'. (Although this scale was described as a 'feelings' questionnaire, a high score related specifically to tiredness, rather than general feelings.)

As recognised by Poffenberger, this scale was "subject to all the errors commonly noted in rating scales, plus those which may be due to the elusive experiences which are to be rated" (p460). Thus, in addition to universal issues such as the relative value of the intervals on the feelings scale and differences between individuals' understanding of the term 'extremely tired', Poffenberger also recognised the importance of scale length and complexity when attempting to measure fatigue - in a population who are, almost by definition, adverse to further effort. This issue in particular continues to concern those attempting to measure fatigue and modern scales remain a compromise between the intricacy of the measurement and the level of intrusion or effort involved. Furthermore, he commented on the difficulty in computing an average change in feelings across individuals where there were considerable differences in the state of feelings before the work begins.

Despite the simplicity of this seven-point one-dimensional rating scale, Poffenberger's findings identified one of the central incongruities facing fatigue research: There appears to be no clear relationship between average loss in performance output, as a result of continuous work, and change in feelings. For each of the four tasks there was a surprisingly uniform loss in 'feeling record', regardless of whether there was a gain or loss in performance over the 5½ hours. As discussed in Chapter One, this finding underlies many of the difficulties associating with measuring fatigue, and highlights the importance of subjective assessment.

Although Poffenberger's work was limited in terms of the development of a *psychometric* scale, the issues raised in this early paper echo throughout the fatigue literature in both the development of appropriate measuring tools and also the advancements in fatigue theory in general. Despite these developments, it was two more decades before there was any more progress in the measurement of subjective fatigue.

4.3.2 A Psychometric Approach

Although the exact nature of the fatigue construct continued to be the subject of much debate during the 1950's, there was a growing agreement that one of the central components of fatigue is a subjective state of tiredness (see Bartley and Chute, 1947). Thus the development of a psychometric scale to reliably measure this state of tiredness was considered to be an appropriate tool for practical purposes and to aid theoretical development.

In the years following Poffenberger, there were several studies which reported employing measures of subjective fatigue such as Barmark (1940), who investigated the effects of caffeine and Foltz, Jung and Cisler (1944), who investigated work output and recovery. However, there were no attempts to develop a scale with psychometric properties until the 1950's when two parallel PhD research programmes were established by McNelly (1954) and Pearson (1957). Both researchers applied Thurstone's scaling technique of equally appearing intervals and each produced two equivalent single scale fatigue questionnaires and subsequently validated them in laboratory studies. While McNelly's aim was to investigate the psychometric properties of a single item verbally anchored rating scale (of the type developed by Poffenberger), Pearson went beyond this, aiming to overcome the criticism of previous questionnaires that the content of the scale was entirely dependent on the assumptions of that particular researcher. His approach followed the procedures set out by Edwards and Kilpatrick (1948) thus, rather than generating his own fatigue terms, items were generated by Air Force personnel who were asked to list words and phrases which might describe the fatigue continuum. These items were then subjected to a series of elimination procedures, which identified items that were ambiguous or not in common usage.

Following this process, 44 items were retained for an experimental version of the checklist, which was used to provide validity estimates and item analyses.

Although both of these research programmes provided some strong support for the argument that fatigue can be reliably measured and that subjective scales can differentiate between different states following fatigue manipulations, this work did little to advance our understanding of the fatigue construct. As Pearson himself acknowledged, he did not address the issue of whether the checklists were one dimensional, simply stating that it was somewhat difficult to give an unequivocal "yes" or "no" to this issue. Furthermore, neither researcher addressed the issue of the dimensionality of the fatigue construct, both assuming that fatigue can be conceptualised and measured as a single entity.

4.3.3 Modern Fatigue Scales

After a short period of quiet, the early 1970's saw a return to the interest in fatigue measurement which has never subsequently subsided. Since this time, a large number of fatigue scales have been reported in the research literature, but there has been a disappointing lack of continuity in this field. The scales have differed widely in their range, dimensionality and focus: Some of the scales have been one dimensional, others multidimensional and the majority have been developed with a narrow focus relating directly to specific research aims. This section will outline one dimensional and multidimensional scales and discuss the contribution of this work to the broader field of fatigue research.

4.3.3.1 One Dimensional Scales

One dimensional fatigue scales have varied greatly in their psychometric sophistication. The most simple one dimensional scales have been single item verbally anchored rating scales with equivalents of "extremely fatigued" at one extreme and "not at all tired" or "full of energy" at the other. Although some authors have endeavoured to investigate the psychometric properties of such scales (e.g. Rhoten, 1982) fatigue research is littered with references to state measurement using such instruments without considering the properties of the scale, e.g. Yoshitake (1971), Rosenfeld, Tenenbaum, Rushkin & Halfons (1989), Neville, Bisson, French and Boll (1994) and numerous others. Beyond these single item scales, a number of more complex one dimensional scales have been developed. These have typically been in the familiar format of multiple item, verbally anchored rating scales; including both stand-alone instruments e.g. The Fatigue Severity Scale (Krupp, LaRocca, Muir-Nash & Steinberg, 1989) and scales which are subcomponents of broader instruments, such as the Fatigue Scale in the Profile Of Mood States (McNair, Lorr & Droppleman, 1971) and the Rand Index of Vitality (Brook, Ware & Davis-Avery, 1979), which formed part of a general mental health battery. In addition to the traditional verbally anchored rating scales, Montgomery (1983) offered an alternative approach to the measurement of a single fatigue dimension in his investigation of 'uncommon tiredness' in undergraduates. His approach was to utilise true/false items to assess the impact of the tiredness on daily functioning. However, in comparison with the more complex one dimensional scales outlined above, instruments such as this offer little in the way of psychometric value and are limited in the extent to which they can differentiate between different levels of fatigue.

Despite the existence of some well-developed scales, one dimensional fatigue instruments have been broadly criticised for their inability to capture the complexity of the subjective experience of fatigue. For example, in an evaluation of existing fatigue scales for use in clinical settings, Smets, Garssen, Bonke and de Haes (1995) noted that individuals with the same overall fatigue score may differ greatly in their experiences: While one person might feel physically exhausted and mentally alert, a second may feel mentally tired, but physically fit. This diagnostic ability of fatigue scales is not only important to health professionals, for both assessment and intervention purposes, and for other applied uses, but also for theoretical advancements of the fatigue construct. However, as previously stated, the vast majority of multidimensional scales have been developed for specific applied uses and have, for the most part, been devised with little reference to the broader literature and without consideration of theoretical issues.

4.3.3.2 Multidimensional Scales

Although the majority of multidimensional fatigue scales have been devised for use in applied health settings, an early exception to this was the work of the Japanese Industrial Fatigue Research Committee during the 1970's. This group delivered a substantial research programme which not only addressed issues of subjective and objective fatigue measurement in industry, but also considered fundamental conceptual issues. Of particular interest to the group was the factor structure of fatigue symptoms and, following an analysis of almost 10,000 workers across various industries, a three dimensional model of fatigue symptoms was proposed (Saito, Kogi & Kashiwagi, 1970). The factors identified were (1) "drowsiness and dullness" or weakened activation, (2) "difficulty in concentration" or decline of work motivation and (3) "projection of physical disintegration" or physical symptoms.

On the basis of this model, two alternative scales were devised; the first was a 30 item checklist which required respondents to state whether or not they were experiencing each of the symptoms - this scale was used by Yoshitake (1978) in his evaluation of patterns of fatigue across different occupational groups. The second instrument was a rating scale which aimed to allow a judgment of fatigue on the basis of a person's appearance (Kashiwagi, 1971). (This was based only on the first and second factors, as the third factor relates more to physical discomfort, such as "*I feel a pain in the waist*", which may be difficult to judge as an observer.) This second scale was particularly interesting, as it was found that judgments of others' appearance have a strikingly similar psychological structure to self-ratings. One seems to be able to discriminate between the dull, listless appearance associated with the dimension of weakened activation and the sluggish, restless appearance associated with weakened motivation. Kashiwagi suggests this is evidence of a link between subjective fatigue feelings and changes in physiological parameters. Furthermore, evidence of this link, coupled with the correlation between subjective self-ratings and observer ratings, supports the general use of a questionnaire method to assess fatigue.

The unusually holistic approach to the fatigue problem adopted by the Japanese Fatigue Committee is essential to unravelling the complex issues posed by this area of investigation. Theoretical advancements were used to guide developments in the assessment and measurement of fatigue, which in turn allowed further theoretical enquiry. However, this work had a disappointing impact on fatigue scales that followed, which is possibly a result of the limited crossover between the work psychology and health psychology disciplines.

4.3.3.3 Fatigue Measurement in Clinical Populations

During the 1980's and 90's a wide range of 'health focused' fatigue questionnaires were developed. These varied broadly in terms of their assumptions about the nature of fatigue, the number of dimensions included and the focus of the measurement: While the emphasis for most was the characteristics and severity of the subjective experience, others focused on the impact of fatigue on daily functioning.

In terms of the type of dimensions included, most scales have incorporated components of mental and physical fatigue, such as the two factor Wessley and Powell (1989) scale and the Chalder Fatigue Scale (Chalder *et al*, 1993), both of which were developed for measuring fatigue in sufferers of Chronic Fatigue Syndrome (CFS). The underlying assumption of the Chalder Scale was that sufferers of Chronic Fatigue experience two discernible types of fatigue; mental and physical. Consequently, items were generated by experts *specifically* to reflect physical and mental symptoms of fatigue - In addition to the circularity problem of such an approach, the content of these subscales remains arguable (despite the inclusion of subject matter experts in the process): In this instrument the physical scale includes items such as "Do you feel sleepy or drowsy?" which could be argued to be sleep-related fatigue rather than physical fatigue. There are also items such as "Are you lacking in energy?", which may be more appropriately conceptualised as general fatigue items. However, although Chalder reports both reliability statistics and Principal Components Analysis (which offer some support of her two factor solution), the development of this scale highlights the lack of coherence in fatigue measurement in general. Although this scale may be valid as a measure of the symptoms in chronic fatigue patients, to classify those high on the 'physical symptoms' scale as physically fatigued in the broader sense may be misleading.

Other scales which incorporate mental and physical dimensions include the FIS: Fatigue Impact Scale (Fisk, Ritvo, Ross, Haase, Marrie, & Schlech, 1994) and the MFI: Multidimensional Fatigue Inventory (Smets *et al*, 1995). The MFI was developed to overcome limitations of previous questionnaires - Smets and colleagues criticised one dimensional measures, for not adequately differentiating between different types of fatigue, and multidimensional instruments, for being too lengthy for completion in patient populations. The exception to this point was the Wessley and Powell (1989) 'Mental and Physical Fatigue Scale', which could be criticised for the opposite problem of being a little too narrow to capture fully the complex subjective experience of fatigue.

The MFI overcomes these criticisms by incorporating five fatigue dimensions in a short scale comprising only 20 items. The five dimensions were postulated to represent the ways in which fatigue can be expressed, such as levels of functioning (General Fatigue); physical sensations related to feelings of tiredness (Physical Fatigue); cognitive symptoms (Mental Fatigue); motivation to start any activity (Reduced Motivation) and reduction in activity (Reduced Activity). Positive and negative items were devised for each subscale, which were then subjected to a fairly rigorous programme of development. This produced evidence of subscale internal consistency and construct validity. The development of this scale represents an excellent attempt to construct a valid and reliable instrument, with sufficient breadth to capture the complexity of subjective fatigue. However, this instrument was developed specifically for clinical populations and was driven by a limited conceptualisation of fatigue (as it relates only to illness).

Like the MFI, the Fatigue Impact Scale (FIS) (Fisk *et al*, 1994) is also a multidimensional scale which includes mental and physical dimensions and the additional 'social fatigue' scale. However, while many other instruments integrate a component of functioning to their scales (e.g. the MAF: Belza, Henke, Yelin, Epstein, & Gilliss, 1993; see below), the FIS scale focuses entirely on the *impact* of fatigue on the daily functioning of patient populations (validated on patients with chronic fatigue, multiple sclerosis, and hypertension). The aim of this scale was to advance understanding of the effects of fatigue on quality of life, which places the instrument even more firmly in the arena of practical health care, rather than the general study of fatigue. Furthermore, close examination of the items raises questions regarding the internal consistency of the subscales: Some items may in fact confound with anxiety and depression domains, such as "I am more moody", "I am more irritable and more easily angered" and "I worry about how I look to other people".

Two further scales of note are the Piper Fatigue Scale (Piper, Lindsey, Dodd, Ferketich, Paul, & Weller, 1989) and the Multidimensional Assessment of Fatigue (MAF) scale (Belza *et al*, 1993). The Piper Scale aimed to measure fatigue in cancer patients along seven dimensions including severity, temporal characteristics and associated symptoms. While this scale has been criticised for being too lengthy and complex for a patient population (Smets *et al*, 1995), the authors clearly state that this instrument was not designed to be used in clinical settings. It was designed to assess *patterns* of fatigue across a variety of health populations, an approach reminiscent of the NASA TLX workload scale (Hart & Staveland, 1988), which evaluates the nature and extent of workload associated with different tasks.

The MAF was derived directly from the Piper Scale to overcome its practical limitations. This scale utilises 16 items to measure four dimensions of severity, distress, degree of life interference and timing. Although this scale meets previous criticisms of practicality, a more recent evaluation of existing fatigue scales brought into question the underlying factor structure of the four conceptual dimensions proposed by Belza (Winstead-Fry, 1998), a criticism which could be levelled at many of the multidimensional scales.

4.3.3.4 Recent Developments in Subjective Fatigue Measurement

Although many of the scales outlined above were validated within particular clinical populations, they were typically designed for evaluating fatigue related to illness *in general*, rather than being designed primarily to measure the unique experiences associated with specific diseases or circumstances. However, during the past decade there has been a growing trend towards more context-specific instruments. This is particularly true in the clinical domain, which has seen a rise in the number of disease-specific fatigue measures. These include the Schwarz Cancer Fatigue Scale (Schwartz & Meek, 1999); the Cancer Fatigue Scale (Okuyama, Akechi, Kugaya, Okamura, Shima, Maruguchi, *et al* 2000) and an instrument to assess the intensity and consequences of fatigue in HIV patients (Barroso & Lynn, 2002). Furthermore, this trend for context specific scales does not appear to be isolated to clinical settings, as evidenced by the Compassion Fatigue scale (Figley, 1995), which was designed for use with practising therapists, and the Multidimensional Fatigue Scale (Matthews, Davies, Westerman & Stammers, 2000) which was designed to measure driving fatigue and includes driving related dimensions of visual fatigue, muscular fatigue, boredom and malaise.

Whilst it is not disputed that these scales may offer sound tools for assessment and diagnosis within specific clinical and occupational populations, the existence of such instruments reflects the lack of generic instruments which have been based on a theoretically driven model of fatigue as a construct, rather than fatigue as a range of disparate symptoms which result from various conditions and diseases.

4.3.3.5 Summary of Existing Fatigue Scales

Although there is little conceptual clarity reflected in existing fatigue scales, there is evidence of some unquestionable trends which provide a firm foundation for future scales and theoretical advancement. The most consistent aspect relates to the multidimensional nature of the construct. In accordance with the findings described in Chapter 3, scale developers have repeatedly been able to separate mental and physical fatigue factors. This again provides support for the notion that these fatigue processes differ on some level. Furthermore, various instruments have psychometrically differentiated between other types of fatigue, e.g. social fatigue, which suggests that a model of fatigue should be more than two dimensional. However, existing scales do not provide an unequivocal picture to the question of how many dimensions there should be, as the resulting factor structures have been dramatically influenced by initial item selection. To complicate the picture further, subscales across instruments, with the same label, often differ dramatically in terms of their actual content. For example, some scales of mental fatigue are characterised by perceived cognitive difficulties (e.g. memory loss etc.) while others define mental fatigue more in energetical terms (e.g. mental alertness).

Finally, existing scales have provided support for the value of the two distinct channels of investigation outlined in Chapter 3. Although no researchers have explicitly labelled their scales as measuring trait or state fatigue, this has been firmly implied in terms of the writing of the items and also the overall instructions. For example, instructions which request respondents to evaluate how they 'currently feel' clearly measure state fatigue, whereas those which assess 'recent' conditions and feelings may be more closely related to trait fatigue. In fact, this review of existing scales reveals a fairly clear trend in the division of these questionnaires, as the more 'work' focused scales have typically been employed to measure subjective state under experimental conditions (e.g. Poffenberger, McNelly, Pearson, Saito and Kashiwagi and the recent Swedish Occupational Fatigue Inventory developed by Ahsberg, Gamberale and Gustafsson (2000) which emphasises physical fatigue following work). On the other hand, health focused measures have been more concerned with evaluating the general status of patient populations and the impact of specific diseases on quality of life (e.g. the Piper scale, the Fatigue Impact Scale, the Multidimensional Fatigue Inventory and the Chalder Fatigue Scale). Therefore, in health settings, the current state of the respondent is of less relevance than in a work setting, were interest is (unfortunately) more likely to be focused on the short-term effects of carrying out specific tasks.

Although there are a few exceptions to this trend, such as the Visual Analogue Scale for Fatigue (Lee, Hicks & Nino-Murcia, 1991) which evaluates state fatigue before and after sleep, these two different approaches highlight the importance of further investigation which explicitly addresses the issues of both state and trait fatigue. Therefore, the remainder of this chapter will describe the development of two instruments; a trait fatigue scale, which continues the 'susceptibility' work embarked upon in Chapter 3, and a state fatigue scale, which emphasises current feelings. Both of these instruments will provide some further insight into the psychometric differentiation between the different types of fatigue and will offer practical solutions to further fatigue research.

4.4 Further investigation of the trait fatigue construct: Development of the 'General Tiredness Questionnaire'

The existing fatigue scales, outlined above, and the findings of the preliminary field work do not offer an unequivocal view of the underlying structure of trait fatigue. However, a number of strong trends have been identified which require further investigation and could provide a basis for the development of a trait fatigue scale, composed of meaningful subscales. This scale could then serve to further our understanding of the nature of the fatigue construct, and act as a practical research instrument.

The factor analysis described in Chapter 3 produced a meaningful six factor solution. However, while this six factor solution appears intrinsically coherent, the subscales require some further development. The main aim of the following psychometric work is to further investigate the factor structure of trait fatigue. This will involve the confirmation of the stability of the six factor solution and the collection of construct and criterion related validity evidence.

4.4.1 Evaluation of trait fatigue factors

Although the six factor solution, detailed in Chapter 3, requires some minor further investigation, the factors identified represent an excellent starting point for this work.

4.4.1.1 Physical (and general) fatigue

A physical fatigue factor was present in all four factor solutions explored. As anticipated, this cluster of items was fairly robust and revealed little overlap with mental fatigue items and sleep-related items. This finding corresponds with the widely held view that mental and physical fatigue differ on some level. However, what was not expected was the inclusion of general fatigue/energy items in the physical fatigue factor for three of the four factor solutions[1]. Although this was unexpected it is not difficult to interpret this finding, as one might expect a strong relationship between general energy levels and inclination to physical work. However, while these dimensions may be closely related, there is some intuitive distinction between general energy and physical fatigue, and therefore, these two components are separated into two distinct factors: physical fatigue and general fatigue. (Although some of the items loading on the general fatigue factor refer to energy/vitality, general fatigue was considered to be a more appropriate factor label as the focus is on trait *fatigue*).

4.4.1.2 Mental fatigue

In the six factor solution, the mental fatigue factor was clearly represented by a cluster of items relating to fatigue following mental activity, concentration and prolonged cognitive work. Although present in the five, seven and eight factor solutions, the clarity of this factor was not replicated. However, mental fatigue is a particularly important component of the scale as the theoretical and practical importance of a mental fatigue factor is undeniable – not only has a large proportion of fatigue research focused on the feelings associated with mental work or mental fatigue as a consequence of illness, but this was the primary source of fatigue as identified in the preliminary interviews (see Chapter Three).

---

[1] This finding provides unexpected support for the inclusion of energy items in the physical dimension of the Chalder Fatigue Scale (Chalder *et al*, 1993)

However, close inspection of the items which load on to this factor revealed two subcomponents of mental fatigue. While some of the items refer to fatigue following mental activity, others refer to the adoption of alternative work strategies when mentally tired. Considering the conceptual distinction between these domains and the importance of strategic adjustment within the theoretical framework of the current research programme, the factor of mental fatigue was separated into two sub factors of mental fatigue and mental strategies.

4.4.1.3 Sleep-related fatigue

The six factor solution produced three sleep-related factors, morning tiredness, evening tiredness and naps. This was again not surprising as few researchers would dispute the proposition that the mechanisms which underlie sleepiness again, on some level, differ from those of mental and physical fatigue. Furthermore there is a general acceptance (partially supported by the preliminary field work) that one can differentiate between the subjective experience of a need to sleep and subjective tiredness following mental work. The individual differences in fatigue at various times of the day (owls and larks) have been widely researched (Horne & Ostberg, 1976) and are attributed to differences in circadian rhythm parameters. Therefore, these differences represent an important component of trait fatigue and hence, it was considered important to include separate dimensions of morning tiredness and evening tiredness in the trait fatigue scale.

An unexpected finding was the very robust factor of naps. The cluster of items referring to nap taking behaviour was consistently clear across all four factor solutions. It appears that individuals possess a tendency to take naps during a day or do not. While this may be interesting, it was considered that this third sleep-related variable would provide little further theoretical value.

4.4.1.4 General fatigue

The first factor in the six factor solution was labelled 'fatiguability'. This incorporated the majority of items which referred to the development of fatigue following the broad range of different circumstances which had been identified by the preliminary field work. As stated in Chapter Three, this factor supports the notion of fatiguability; that some individuals are more prone to developing fatigue when they interact with the world. Although this is an interesting finding, the current analysis is interested in the underlying clusters of fatiguing conditions and the nature of their relationships. Furthermore, it would be difficult to identify a meaningful set of items from this factor which could reliably represent 'general' fatiguing conditions. Therefore, this general fatiguability factor was not included in further scale development and was replaced by the alternative conception of general fatigue identified within the physical factor.

4.4.1.5 Emotional fatigue

One final dimension of the fatigue construct worthy of consideration is emotional fatigue. Although the preliminary field work and existing research suggested that there may be a separate factor of emotional fatigue, none of the four factor solutions separated this cluster of items into a coherent factor. In each analysis, the items relating to emotional problems or difficulties loaded onto either general fatiguability or mental fatigue. This is further supported by Aisbitt (1998) who, in a similar investigation of trait fatigue, found that mental and emotional fatigue items consistently loaded onto the same factor.

This suggests that the processes which underlie emotional fatigue do not differ from those underlying mental fatigue. On further consideration, this is intuitively sensible as dealing with emotional problems requires a component of problem solving, and therefore emotional 'work' may involve the same cognitive mechanisms as those involved with mental work. Again this is a theoretically interesting proposition. However, there is currently no support for including a factor of emotional fatigue as this domain failed to load as a separate factor. Therefore, emotional fatigue items were not included in further analyses.

4.4.2 Factor components of the General Tiredness Questionnaire

Considering various theoretical perspectives, existing fatigue scales and the analysis of the various factor solutions outlined in Chapter 3, a six factor trait fatigue scale was developed which was entitled the General Tiredness Questionnaire. This incorporated the following factors; *General Tiredness, Mental Fatigue, Mental Strategies, Physical Fatigue, Morning Tiredness* and *Evening Tiredness.* Selection of the items to represent each factor followed the procedures set out by Kline (1993), and hence, item inclusion was based on factor loadings and item mean and variance. As a result of rigorous item writing, none of the 98 items had to be excluded on the basis of an extreme mean or low variance. On a six point scale, item means ranged between 2.51 (*I come across to others as a bit lethargic*) and 4.67 (*I feel more lively when out of doors*) and only one item had a standard deviation below 1.0 (*spending time with friends is invigorating*). Factor loadings were considered for the Physical Fatigue, Morning Tiredness and Evening Tiredness factors. However, this index was much less relevant for the adapted factors: Mental Fatigue, Mental Strategies and General Fatigue, as these are adapted scales and therefore loadings no longer relate directly to the items.

Four items were selected to represent each of the six factors - with the exception of the Physical Fatigue factor, as only three items reached the selection criteria for this subscale. Therefore, an additional item was written to represent this factor. Consequently, the new multidimensional trait fatigue questionnaire comprised 24 items, representing six fatigue dimensions. (A copy of this questionnaire can be found in Appendix 3). A further amendment from the first development scale was the conversion to a 5 point scale, as a 'neither agree nor disagree' option was requested by a small number of respondents from the first sample. This seemed to be reasonable, as it is clearly possible to best describe oneself as average with regard to these characteristics.

4.4.3 Second development sample

To evaluate the proposed factor structure, a second development sample was tested. Two hundred and ninety nine questionnaires were completed and returned from four hundred and eighty seven distributed, which represented a response rate of 61%. This sample was composed entirely of students due to the availability of this group. It consisted of 199 (66.5%) females and 86 (28.7%) males (14 respondents did not state their gender). The age range of the sample was 18 to 43 (mean = 21.4; sd = 4.4).

As with the first development sample, outlined in Chapter 3, the data were coded, entered into SPSS and subjected to data reduction. Principal components analysis with varimax rotation was used and eigenvalues greater than one were accepted. Six factors had eigenvalues greater than one, which together accounted for 60.6% of the variance (see Appendix 4 for factor solution). Only factor loadings above .3 were acknowledged.

4.4.3.1 Analysis of the six factor solution

The first factor, ***Physical Fatigue***, accounted for 14.2% of the variance. All four physical fatigue items loaded strongly onto this factor (factor loadings ranged from .708 to .799). In addition to the four physical items, two general fatigue items also loaded on this factor; "*I am not a lively person*" possessed a loading of .470 and "*I think of myself as having bags of energy*" possessed a factor loading of -.310. This reinforces the relationship between physical fatigue and general fatigue identified in the initial factor analysis. However, both of these factors possessed higher loadings on the General Fatigue factor, and hence were retained as General Fatigue items.

The second factor, ***General Fatigue***, accounted for 13.3% of the variance. All four general fatigue items loaded onto this factor (factor loadings ranged from -.571 to .634). In addition to the general items, two of the morning tiredness items also loaded strongly onto this factor; "*I wake up feeling tired*" possessed a loading of .652 and "*However tired I am I feel refreshed after sleep*" possessed a loading of -.662. These items are clearly consistent with general fatigue as well as morning tiredness because those people with generally low levels of energy may be more likely to wake up feeling tired. Surprisingly, these items failed to load onto the morning tiredness factor above the cut-off level of .3. Therefore, these two items were removed from further analysis.

Six items loaded onto the third factor of ***Mental Fatigue*** (accounting for 10.5% of the variance). In addition to the 4 mental fatigue items (factor loadings ranging from .549 to .817) one of the mental strategies items ("*If I am feeling tired during work I try to do less* *demanding tasks*") and one of the general items ("*I feel alert most of the time*") also loaded onto Mental Fatigue, although these items were found to also load onto their target factors. Therefore, these items were retained to represent their original factor.

The ***Evening Tiredness*** factor accounted for 8.12% of the variance. This factor comprised three of the evening tiredness items (factor loadings ranging from .373 to .784) and also one of the mental fatigue items "*Prolonged mental activity wears me out*". However, this item possessed a higher loading onto mental fatigue and was therefore retained as a component of the mental fatigue subscale. The item selected to represent negative evening tiredness ("*I feel at my most alert later in the day*") did not load onto this factor (above .3) but positively loaded onto the morning tiredness factor (loading = .787). This is not surprising as morning tiredness and evening tiredness would be expected to correlate negatively. However, in the light of the current factor analysis, this item was considered to be more appropriate as a component of morning tiredness.

The ***Mental Strategies*** factor was much clearer, as only the four items selected to represent mental strategies loaded onto this factor. Factor loadings ranged from .365 to .787 and the factor accounted for 7.98% of the variance.

The ***Morning Tiredness*** factor (accounting for 6.33% of the variance) was composed of only two of the morning tiredness items (loadings of .321 and .490) in addition to the item originally selected to represent evening tiredness. Hence, three items were selected to represent this factor.

Therefore, the General Tiredness Questionnaire comprised 22 items, four items each for Physical Fatigue (items 4, 9, 13 & 19 in Appendix 3), General Fatigue (items 1, 6 reversed, 12 & 17 reversed), Mental Fatigue (items 7, 10, 22 and 23 reversed) and

Mental Strategies (items 15, 16,18, 24), and three items each for Evening Tiredness (3, 5, 11) and Morning Tiredness (8, 20, 21). (Items 2 & 14 in Appendix 3 were removed from further analysis.)

4.4.4 Validation of the General Tiredness Questionnaire

4.4.4.1 Measures

A composite questionnaire entitled 'General Feelings Questionnaire' was produced, which contained a range of psychometric scales selected to assess construct validity (evidence which demonstrates that the instrument relates to the particular construct it is supposed to: Kline, 1993) and criterion related validity (measures of the relationship between scores on the instrument and external criteria). Two forms of construct validity were collected: convergent (relationships with other scales proposing to measure the same construct) and divergent (relationships with scales proposing to measure conceptually distinct constructs with hypothesised relationships). In addition to the 22 item General Tiredness Questionnaire, the Multidimensional Fatigue Inventory: **MFI** (Smets *et al*, 1995) was included to measure convergent validity. This instrument comprises five subscales, namely *general fatigue, physical fatigue, reduced activation, reduced motivation* and *mental fatigue*. Three further psychometric instruments were included to measure divergent validity. These were

- the Mental Toughness Questionnaire: **MTQ 48** (Clough, Earle & Sewell, 2002)
- the State Trait Anxiety Inventory: **STAI Y-2** (Spielberger, 1983)
- and the Hospital Anxiety and Depression Scale: **HAD** (Zigmond & Snaith, 1983)

Also included was an introductory page, including generic instructions and personal and job-related information for criterion related validation. This focused on self and family life and the respondent's job, in terms of the nature of the work demands and opportunities for control, support and coping. (A copy of this composite questionnaire is available in Appendix 5)

4.4.4.2 Sample

One hundred questionnaires were distributed within four organisations. Eighty five of these were returned, yielding an 85% response rate. The organisations were a leading bank (n=22), teachers from a high school (n=25), a printing firm (n=22) and students from Hull University (n=16). A diverse sample was actively sought to ensure some variability in the range of working environments. The sample was composed of 40 males and 45 females, the age range was 18-58 and the mean age was 37.76 (sd =11.87). With regard to marital status, 46 were married, 29 were single, 10 were divorced and no participants were recorded as widowed or separated. 63% of the participants had children.

4.4.4.3 Data Treatment

The data from each questionnaire were reduced into their relevant subscales, as outlined in Table 4.1.

GTQ	MFI	STAI Y-2	MTQ 48	HAD
General Fatigue Mental Fatigue Physical Fatigue Morning Tiredness Evening Tiredness Mental Strategies	General Fatigue Physical Fatigue Mental Fatigue Reduced Activity Reduced Motivation	Trait Anxiety	Challenge Commitment Control (of life) Control (of emotions) Confidence (in abilities) Confidence (interpersonal)	Anxiety Depression

Table 4.1: Subscale breakdown for validation questionnaires

To investigate the convergent validity of the General Tiredness Questionnaire, responses were correlated with those from the MFI. Divergent validity was evaluated on the basis of correlations between the General Tiredness Questionnaire and the STAI Y-2, MTQ 48 and the HAD scale.

Criterion-related validity was assessed using multiple regression analyses to investigate the prediction of trait fatigue from life and job factors.

4.4.4.4 Results

Descriptive Data for the General Tiredness Questionnaire

Table 4.2 provides the descriptive statistics for the six trait fatigue factors, for the sample as a whole and for males and females. For the general sample, all mean scores are close to the midpoint (3), with the exception of evening tiredness and morning tiredness. The mean score for evening tiredness is fairly high (3.75), whereas, the mean score for morning tiredness is fairly low (2.37). This suggests that, in general, individuals are more prone to feeling tired in the evening, than in the morning. Interestingly, there is no difference between morning tiredness in males (2.36) and females (2.38), but there is a significant difference between males (3.55) and females (3.93) for evening tiredness (t = 2.47; df = 83; $p < 0.05$). This finding concurs with established views regarding gender differences in circadian rhythms, as males are more frequently categorised as 'owls' (Horne & Ostberg, 1976).

	Minimum	Maximum	Mean	S.D.	Gender	Mean	S.D.
GTQ general fatigue	1.00	4.25	2.61	0.80	male	2.49	0.84
					female	2.71	0.76
GTQ morning tiredness	1.00	4.67	2.37	0.70	male	2.36	0.63
					female	2.38	0.76
GTQ evening tiredness	1.33	5.00	3.75	0.73	male	3.55	0.76
					female	3.93	0.65
GTQ physical fatigue	1.00	5.00	3.14	0.97	male	2.92	0.96
					female	3.34	0.94
GTQ mental fatigue	1.25	4.50	3.09	0.61	male	3.06	0.61
					female	3.12	0.63
GTQ mental strategies	1.25	4.50	3.17	0.71	male	3.15	0.76
					female	3.19	0.68

Table 4.2: Descriptive statistics for trait fatigue factors - by gender

Other gender differences are found for the GTQ factors of general fatigue (Males, 2.49; Females, 2.71) and physical fatigue (Males, 2.92; Females, 3.34), but only the gender differences for physical fatigue reached statistical significance ($t = 2.02$; df = 83; $p < 0.05$).

Scale Reliability

Cronbach's alpha was calculated for each sub scale as a measure of internal consistency (see Appendix 6). With the exception of the Morning Tiredness factor (0.76), all subscales reached the acceptable level (0.80) recommended by Kline (1993). This largely supports the homogeneity of each subscale.

Construct Validity

General Tiredness Questionnaire and the MFI: To investigate the convergent validity of the GTQ, Pearson's correlations were calculated for the six GTQ factors and the five factors of the MFI (see Table 4.3). Moderate to strong correlations were expected between the subscales of the two trait fatigue instruments. For instance, the two general fatigue scales should be positively correlated, as should the two physical scales and the two mental scales. However, while the General Tiredness Questionnaire aims to measure the predisposition to develop fatigue following different experiences, the MFI aims to measure how respondents have been feeling lately. Therefore, the measures of trait fatigue in the MFI are likely be confounded by recent activities. So, although there was an expectation that certain factors of the two scales would be positively correlated, they are based upon different approaches to trait measurement.

Despite these differences, there was a strong positive correlation between the two factors of general fatigue ($r = 0.71$, $p < 0.01$) and a moderate correlation between the two factors of physical fatigue ($r = 0.49$, $p < 0.05$). However, the relationship between the GTQ general factor and the MFI physical factor ($r = 0.67$, $p < 0.01$) was surprisingly stronger than the two physical factors. This confirmed earlier findings which highlighted a strong relationship between general energy and physical fatigue.

	GTQ General	GTQ Morning	GTQ Evening	GTQ Physical	GTQ Mental	GTQ Strategies	MFI General	MFI Physical	MFI Reduc. activation	MFI Reduc. motivation
GTQ Morning	**.30**									
GTQ Evening	**.34**	.00								
GTQ Physical	**.39**	**.26**	**.21**							
GTQ Mental	.16	.18	**.36**	.02						
GTQ Strategies	.15	**.28**	.17	**.20**	**.25**					
MFI General	**.71**	**.37**	**.40**	**.42**	**.27**	.10				
MFI Physical	**.67**	**.38**	**.28**	**.49**	**.21**	.15	**.83**			
MFI Red. Activ.	**.58**	**.39**	.13	.41	.15	**.25**	**.56**	**.62**		
MFI Red. Motiv	**.57**	**.33**	**.23**	**.24**	**.30**	**.22**	**.57**	**.53**	**.61**	
MFI Mental	**.38**	**.35**	.08	.15	**.39**	.17	**.48**	**.37**	**.40**	**.60**

Table 4.3: Correlation matrix for General Tiredness Questionnaire (GTQ) factors and Multiple Fatigue Inventory (MFI) factors (significant relationships in bold; $p < 0.05$)

This relationship was further confirmed by the very strong correlation between the MFI physical and general scales ($r = 0.83$, $p < 0.001$) which supports the proposition that the mechanisms underlying these two types of fatigue are closely related. While this is consistent with the current model of fatigue, the level of relationship between physical and general fatigue may not be as strong as the MFI suggests, because the correlation between the general and physical factors in the GTQ is not nearly as strong ($r = 0.39$, $p < 0.05$). One possible explanation for this lies in the different approaches taken to measure trait fatigue: The MFI asks respondents to consider "how you have been feeling lately". Respondents are then presented with a list of statements such as "I feel fit" and "Physically I feel in a bad condition", which could be described as assessing recent states, rather than fatigue predisposition. Although the approach taken by the MFI is not unusual (e.g. STAI), there are two possible limitations to it. First, it is likely that 'general' feelings are confounded by recent activities. For example if an individual has recently undertaken heavy physical work, they may be more inclined to say that they have recently been physically tired, irrespective of their predisposition to trait physical fatigue. Secondly, it could be argued that this approach differs from the approach taken by the General Tiredness Questionnaire, as an individual responding to the MFI may evaluate recent feelings by first focusing on their current state and then comparing this to previous days/weeks. Whereas, the GTQ was designed to evaluate the extent to which individuals are likely to develop fatigue following physical work, for example. This requires the individual to consider a time that they have carried out heavy physical work and recollect the consequences. As intended, this latter approach may be less contaminated by current state. The data presented in Table 4.2 are consistent with this hypothesis. In general, the internal relationships between the five MFI scales are much higher than the relationships between the six GTQ scales. For the MFI, the relationships between the sub scales range from weak ($r = .37$, $p < 0.05$) to very strong relationships ($r = .85$, $p < 0.05$). These are substantially higher than the GTQ sub scales, the strongest relationship being only a weak correlation ($r = .39$, $p < 0.05$). This suggests that the subscales of the GTQ are more independent than those of the MFI.

General Tiredness Questionnaire and Anxiety and Depression: To investigate the divergent validity of the GTQ, Pearson's correlations were calculated for the six GTQ factors and the measures of anxiety and depression (STAI Y-2 & HAD). These correlations are presented in Table 4.4. As one would expect, the divergent relationships between the measures of anxiety and depression and the GTQ scales were less strong than the convergent relationships outlined above. The relationship between anxiety and fatigue is well-documented (see Hockey, 1993 – who argues both constructs are components of 'strain'). As expected Table 4.4 revealed a moderate correlation between GTQ general fatigue factor and the STAI Y-2 ($r = .41$, $p < 0.05$). Furthermore, there is a significant correlation between the GTQ general factor and the HAD depression scale ($r = 0.36$, $p < 0.05$).

Consideration of Table 4.4 highlights further relationships of interest between the GTQ mental and physical scales and the validation measures. While the GTQ physical scale is significantly correlated with depression, but not anxiety, GTQ mental fatigue is significantly correlated with anxiety, but not depression. While these differences are only small, this provides further good evidence of the divergent validity of the dimensions; as depression is characterised by low energy (hence a relationship with physical fatigue), while anxiety may be considered as a form of mental workload resulting from the cognitive requirement to problem solve, which may in turn generate mental fatigue.

Finally, further evidence of divergent validity is provided by the absence of a relationship between the GTQ strategies scale and either of the validation measures. One would not expect a relationship between the tendency to use strategies for dealing with mental fatigue and anxiety or depression.

	GTQ General	GTQ Morning	GTQ Evening	GTQ Physical	GTQ Mental	GTQ Strategies
STAI Y-2	**.41**	**.22**	**.31**	.13	**.25**	.11
HAD Anxiety	**.26**	.05	**.24**	.13	**.26**	-.05
HAD Depression	**.36**	.20	**.20**	**.29**	.20	.18

Table 4.4: Correlation table for trait fatigue factors, STAI and HAD (sig. relationships in bold, $p < 0.05$)

General Tiredness Questionnaire and Mental Toughness: In addition to the evidence of divergent validity provided by the anxiety and depression measures, the relationships between the subscales of the GTQ and the Mental Toughness Questionnaire (MTQ) were also investigated. Table 4.5 presents the correlations between the six GTQ scales and the six MTQ scales; *challenge, commitment, control (of life), control (of emotions), confidence (in abilities)* and *confidence (interpersonally).* As mental toughness is defined as resilience and self efficacy, one would expect moderate negative relationships with fatigue. As anticipated, GTQ general fatigue and the MTQ scales were all negatively correlated, two scales moderately so ($r = -0.43$ to $r = -0.40$, $p < 0.05$) and a weak negative correlation was found with the remaining scales ($r = -0.39$ to $r = -0.24$, $p < 0.05$). Furthermore, the general pattern of relationships between the two sets of subscales provides additional evidence of construct validity. For example GTQ mental fatigue is moderately negatively correlated with MTQ commitment, suggesting that individuals with a predisposition to trait mental fatigue may find it difficult to complete tasks they begin.

	GTQ General	GTQ Morning	GTQ Evening	GTQ Physical	GTQ Mental	GTQ Strategies
MTQ Challenge	**-.40**	-.18	-.17	-.20	-.08	-.13
MTQ Commitment	**-.43**	**-.38**	-.20	**-.25**	**-.33**	**-.24**
MTQ Control/emotion	**-.32**	**-.25**	-.18	**-.24**	-.18	**-.24**
MTQ Control/life	**-.39**	**-.26**	**-.32**	-.09	-.15	-.12
MTQ Confidence/ability	**-.31**	-.14	-.20	-.04	**-.24**	.01
MTQ Confidence/interpersonal	**-.24**	-.03	-.15	**-.22**	-.05	.12

Table 4.5: Correlation table for GTQ factors and MTQ factors (sig. relationships in bold; $p < 0.05$)

Construct Validity Summary: The correlations between the GTQ fatigue factors and the anxiety, depression and mental toughness measures provide good evidence for the construct validity of the GTQ as a measure of trait fatigue. In terms of convergent validity, the instrument correlates moderately to strongly with the alternative measure of trait fatigue (MFI), while improving on this instrument by maintaining low relationships between its own subscales. This suggests that the GTQ subscales are, in fact, measuring what they claim to and that there is less overlap between the scales than exists within the MFI. Evidence of divergent construct validity is provided by relationship with other related constructs, anxiety, depression and mental toughness. All relationships were in the direction expected, but the low to moderate relationships with the GTQ subscales indicate that the construct of trait fatigue is clearly distinct from those of anxiety, depression and mental toughness.

Criterion Related Validity

The criterion related validity of the GTQ subscales was investigated using the data collected on the introductory page of the composite questionnaire entitled the General Feelings Questionnaire. This included information relating to work, in terms of demands and opportunities, and life, in terms of the individuals' family situation.

The prediction of GTQ fatigue factors from work and life circumstances is not straightforward. For example, while one might expect recent mental demands to predict trait mental fatigue as conceptualised by the MFI, mental demands would not be expected to predict individual predisposition to developing fatigue following mental work. This rationale is also true for GTQ physical fatigue factor. However, counter-intuitively, work demands are more likely to be significant predictors of evening tiredness, as this factor includes questions such as *"I feel tired in the evenings after work"*. Therefore, a series of Multiple Regression analyses were carried out on the GTQ subscales with carefully targeted IVs, which could possess a meaningful relationship with the factors. As the aim of these analyses was simply to identify which, if any, of the IVs were significant predictors of the GTQ factors, the *Standard* Multiple Regression method was employed as recommended in Tabachnick & Fidell (1996)

G*TQ General Fatigue:* As the GTQ subscale of general fatigue aims to measure predisposition to general tiredness, ability to cope with demands and number of children were the only criteria which could be expected to predict GTQ general fatigue. An average of emotional, physical and mental demands was calculated and included in the analysis to confirm that trait fatigue is not a consequence of recent demands. The results of this analysis are presented in Table 4.6. As expected, ability to cope (*beta* = -

.31, $p < 0.01$) and number of children (*beta* = .24, $p < 0.05$) were both found to be significant predictors. Furthermore, recent work demands did not predict trait general fatigue, providing some support for the hypothesis that the General Tiredness Questionnaire does not simply measure recent states, but relatively stable characteristics.

	GTQ General	**Coping**	**Children**	**Mean Demands**	**B**	***B***	**t**	**Sig.**
Coping	-.35				-.40	-.31	-3.10	p<0.01
No. Children	.29	-.15			.19	.24	2.37	p<0.05
Mean Demands	.02	-.10	.09		-.02	-.02	-0.27	NS
mean	2.61	3.21	1.15	3.14				
sd	0.80	0.64	1.02	0.76				R = .42 R^2 = .18 Adj R^2 = .15

Table 4.6: Multiple Regression summary table: Predictors of GTQ General Fatigue

GTQ Physical Fatigue: The GTQ physical fatigue scale aims to measure predisposition to developing fatigue following physical demands. Therefore, one would expect age and sex to be significant predictors of this factor. The rating of job-related physical demands was also included in this analysis; although this was not expected to contribute to the variance in trait physical fatigue. The results of this analysis (presented in Table 4.7) were again as expected. Age and sex were both found to be significant predictors of trait physical fatigue. Older respondents were generally found to have higher levels of trait physical fatigue (*beta* = .24, $p < 0.05$) and women were also found to have generally higher levels of trait physical fatigue than men (*beta* = .21, $p < 0.05$). Furthermore, recent physical demands at work did not predict trait physical fatigue.

	GTQ Physical	Age	Sex	Phys Dem	B	β	t	Sig.
Age	.24				.01	.24	2.28	$p<0.05$
Sex	.21	.03			.41	.21	2.00	$p<0.05$
Physical Demands	.03	.14	.15		-.02	-.03	-0.28	NS
Mean	3.14	37.76	1.53	2.65				
Sd	0.97	11.87	0.50	1.18				R = .32
								$R^2 = .10$
								Adj $R^2 = -.07$

Table 4.7: Multiple Regression summary table: Predictors of GTQ Physical Fatigue

GTQ Evening Tiredness: Evening tiredness differs slightly from the above two factors, in that it is more likely to confound predisposition to evening fatigue with recent events. Unlike the general, physical and mental factors, asking respondents about how tired they are in the evenings may lead them to focus upon recent evenings. Because of this, measurement of this factor may be confounded with recent states, work demands and opportunities. Therefore, a broader range of criteria were included in this analysis to investigate this proposition. However, the data presented in Table 4.8 suggest that the measurement of this factor is not confounded with recent conditions/events as only age (*beta* = .21, $p < 0.05$) and sex (*beta* = .23, $p < 0.05$) were found to be significant predictors of trait Evening Tiredness. This is consistent with current perspectives on circadian rhythms, as women might be expected to have higher trait evening fatigue than men. However, the relationship between age and evening tiredness is more complicated as it may be curvilinear; for example it is possible that evening tiredness increases up to a certain age, and then possibly decreases in retirement age, as the restorative function of sleep becomes less crucial. However, this is merely supposition and, in the current sample, one would expect age and evening tiredness to be approximately linearly related, which provides an explanation for the results presented here.

	GTQ Evening	Cop	Age	Sex	Ment dems	Physi dems	Emotl Dems	Control	Support	B	β	t	Sig.
Coping	-.18									-.08	-.07	-0.75	NS
Age	.32	-.14								.01	.21	2.07	$p<0.05$
Sex	.26	-.09	.03							.34	.23	2.22	$p<0.05$
Mental Demands	.19	.10	.19	-.02						.12	.16	1.21	NS
Physical Demands	.04	-.09	.14	.15	-.26					-.05	-.09	-.82	NS
Emotional Demands	.34	-.17	.24	.3	.45	.21				.09	.16	1.25	NS
Personal Control	-.24	.14	-.15	-.03	.29	-.18	-.01			-.12	-.20	-1.89	NS
Personal Support	-.29	.13	-.01	.19	.23	-.18	.07	.31		-.11	-.13	-1.26	NS
mean	3.75	3.21	37.76	1.53	3.82	2.65	2.96	3.11	3.22				
sd	0.73	0.64	11.87	0.50	0.97	1.18	1.31	1.21	0.89				R = .54
													R^2 = .29
													Adj R^2 = .21

Table 4.8: Multiple Regression summary table: Predictors of GTQ Evening Tiredness

GTQ Morning Tiredness: Surprisingly, the Multiple Regression analysis presented in Table 4.9 reveals that neither age, sex nor coping significantly predict GTQ morning tiredness, despite the possible relationships between morning tiredness and age and sex. Therefore, to further investigate this finding a correlation matrix including all criterion variables was produced, but no significant correlations were revealed. It would appear that trait morning tiredness is unrelated to work demands and life circumstances, which is particularly surprising as GTQ evening tiredness does correlate with these variables. This analysis provides no clear explanation for these findings.

	GTQ Morning	Coping	Age	Sex	B	*β*	t	Sig.
Coping	-.03				-.06	-.05	-0.53	NS
Age	-.14	-.14			.09	-.15	-1.39	NS
Sex	.01	-.09	.03		.01	.01	0.11	NS
Mean	2.37	3.21	37.76	1.53				
Sd	0.70	0.64	11.87	0.50				R = .15 R^2 = .02 Adj R^2 = -.01

Table 4.9: Multiple Regression summary table: Predictors of GTQ Morning Tiredness

GTQ Mental Strategies: Table 4.10 presents a summary of the Multiple Regression analysis of GTQ mental strategies. Coping, personal control and personal support were included in this analysis as it was hypothesised that they may predict the use of short cut strategies when mentally fatigued. However, none of these variables was found to predict mental strategies. Again, the explanation for this is unclear.

	GTQ Strategies	Coping	Control	Support	B	*β*	t	Sig.
Coping	-.04				-.04	-.03	-0.32	NS
Control	.06	.14			.07	.13	1.13	NS
Support	-.16	.13	3.12		-.15	-.19	-1.71	NS
Mean	3.17	3.21	3.11	3.22				
Sd	0.71	0.64	1.21	0.89				R = .20 R^2 = .04 Adj R^2 = -.00

Table 4.10: Multiple Regression summary table: Predictors of GTQ Mental Strategies

GTQ Mental Fatigue: Finally, the possible evidence for GTQ mental fatigue was considered. However, in the current validation study, it was hypothesised that none of the current criterion variables should predict the predisposition to mental fatigue. While small but significant correlations were found with mental demands (r = .26, $p < 0.01$) and emotional demands (r = .21, $p < 0.01$), none of these variables was found to be significant predictors of GTQ mental fatigue in the multiple regression analyses (therefore, no regression table is included). It is hypothesised that these weak correlations are the result of confounding recent circumstances, rather than evidence of criterion validity of this subscale.

4.4.4.5 Summary of General Tiredness Questionnaire Validation

While the validation study of the GTQ provided strong evidence of construct validity (in terms of both its convergent and divergent relationship with other validated scales), the criterion related evidence for the GTQ was mixed. Although some good evidence was found in support of the GTQ general fatigue, physical fatigue and evening tiredness scales, there was no criterion related evidence to support the GTQ mental fatigue, mental strategies and morning tiredness scales.

It remains unclear whether this lack of evidence reflects limitations inherent in these scales or within the criteria. However, the substantial body of construct validity evidence and the robustness of the factor analysis is certainly sufficient evidence to support the proposition of trait fatigue as a multidimensional construct. Furthermore, the evidence presented earlier in the chapter supports the approach of assessing trait fatigue via predisposition, rather than recent states.

4.5 Further investigation of the state fatigue construct: Development of the 'Feelings Questionnaire'

Having carried out further analysis of the psychometric properties of trait fatigue, a similar process was undertaken with regard to state fatigue. As with the trait fatigue scale, the aim was primarily to investigate the factor structure of state fatigue, but there was an additional aim of developing a research instrument capable of measuring the different state fatigue components pertinent to a series of experiments investigating responses to mental and physical work. The preliminary fieldwork, existing fatigue scales and theoretical perspectives together provide a useful starting point for this work.

4.5.1 Preliminary field research and existing questionnaires

As outlined earlier in the chapter, the measurement of subjective state fatigue is problematic. While feelings of tiredness are extremely commonplace, the experience is difficult to define. This was highlighted by the preliminary interviews (described in Chapter 3) which identified that, while some individuals do recognise some differences in the feelings of fatigue generated by different circumstances, describing the nature of these differences is more difficult. Nonetheless, the Critical Incident interviews did provide some support for a differentiation between sleepiness, mental and physical tiredness and, to a limited extent, emotional fatigue. Thus, differentiation between these types of fatigue may be useful in understanding how fatigue develops and the relative changes across the different dimensions. However, in the earlier investigation into the structure of trait fatigue, factor analysis consistently failed to identify an emotional fatigue factor; the items relating to emotional problems or difficulties loaded onto either general fatigability or mental fatigue. This suggests that the processes which underlie emotional fatigue may not differ from those underlying mental fatigue. While it is likely that the structure of state fatigue may differ from that of trait fatigue, it could be argued that it is in fact less likely to be multidimensional. This issue will be returned to in Chapter Ten. However, in the absence of any supporting evidence, no items of emotional fatigue were included in the state fatigue scale.

Thus, the subjective data collected during the interview phase supported state dimensions of *mental*, *physical* and *sleep-related* fatigue. In addition to this, existing fatigue questionnaires also provided a foundation for the development of the current scale. As previously discussed, a broad range of fatigue questionnaires have been developed, which have varied both in the number of dimensions considered and their measurement focus: evaluation of current state or general susceptibility to tiredness. State fatigue has most frequently been measured using one-dimensional questionnaires to assess the extent of current tiredness. But, while such scales clearly have some useful applications, they could be argued to be insufficient to capture the full complexity of the subjective experience of fatigue.

Few multidimensional state fatigue scales exist and (in common with trait fatigue scales) there is some variability with regard to the dimensions. However, the themes within state measurement are more consistent and less broad: As previously discussed, they typically differentiate between physical and mental fatigue and often possess items relating to sleepiness or sleep-related fatigue – although this factor label is not frequently made explicit. The most notable state scale development work was that by the Japanese Industrial Fatigue Research Committee in the 1970's (see section 4.3.3.2). The three factors incorporated within their 30 item checklist were; difficulty in concentration, projection of physical disintegration and drowsiness. These three subscales essentially represent mental, physical and sleep-related state fatigue, and are not only recurring themes throughout the fatigue literature, but are consistent with the findings from the preliminary interviews.

4.5.2 Scale Construction

Thus, unlike the inconsistent picture of trait fatigue which has evolved from the research literature, there is a more consistent picture of the taxonomy of state fatigue (although it could be argued that this may be partially a reflection of its rarity). However, as existing dimensions correspond to the types of fatigue which were generated from the preliminary work, and they are pertinent to the current research programme, the development of the state scale was focused on the dimensions of mental, physical and sleep-related fatigue.

The requirements for the state scale were different from those of the General Tiredness Scale. As originally identified by Poffenberger (1928) scale length is of particular importance in the self report of a population who are fatigued. Therefore, the number of scale items must be kept to a minimum and items should not be unnecessarily long or complex. Considering these constraints, a fairly narrow range of short questionnaire items was generated, with the intention of adequately representing the domains of mental, physical and sleep-related fatigue. Twelve items were produced which again aimed to follow Kline's (1986) guidelines for writing clear, unambiguous and specific items. While the standard approach to scale development is to generate subscales of equivalent length, this approach was not adopted here. In the measurement of state fatigue, a key component of questionnaire design is to keep the length of the questionnaire to a minimum. Therefore, items were written and selected with the aim of adequately representing each domain. This resulted in an unequal number of questions, as the different types of fatigue are differentially complex (see Table 4.11)

These 12 items were then piloted with 10 respondents, who were asked to assess clarity and ambiguity. All items were considered to be suitable by all respondents. However, following discussion with a subject matter expect, item 8 "I feel physically strong" was removed from further analysis as it was argued that this may be more representative of general health and fitness rather than low state physical fatigue. Thus, eleven state fatigue items remained.

4.5.2.1 State fatigue and related constructs: Boredom and negative affect

Having identified the importance of mental, physical and sleep-related fatigue and provided some support for distinguishing between these factors, two further dimensions of subjective state were considered. To investigate fully the development of fatigue in any given circumstances, one must also consider the two distinct but fatigue-related states of boredom and negative affect. An important conceptual issue in the development of the fatigue construct is whether fatigue can be differentiated from other cognitive states such as boredom, stress, anxiety and motivation. This is particularly relevant within the context of the adopted theoretical model: In their 1998 paper, "The construct of psychological fatigue", Hockey and Meijman argue that fatigue may normally be distinguished from both boredom and negative affect: In the case of boredom, fatigue can be distinguished in terms of its goal management properties, i.e. the presence or absence of engagement with goal directed behaviour: While fatigue may develop as a result of prolonged task engagement or high workload, boredom conversely is associated with underload and task disengagement – although feelings of sleepiness may be common to both states, aversion to further effort is only hypothesised to be a feature of fatigue. In the case of anxiety, fatigue can be distinguished in terms of the role of negative affect in state regulation: Unlike fatigue, negative affect can be seen as a product of a mismatch between current and goal states. Hence, negative affect may occur following the decision to adopt a demanding goal, while fatigue may occur only as a result of carrying out the behaviour directed at meeting an adopted goal.

Therefore, because of the close relationship between these states and fatigue, it was considered important to incorporate dimensions of negative affect and boredom in the current state fatigue scale. To represent boredom, two new items were included; "I feel bored" and "I feel detached/uninterested". To represent negative affect, three new items were included: “I feel uneasy”, “I feel tense/on edge” and “I feel irritated and annoyed”. These five new items were then incorporated into the three factor state fatigue scale, resulting in a 16 item instrument (see Table 4.11).

1. I feel mentally tired	Mental Fatigue
2. I don’t feel like making much of an effort	Mental Fatigue
3. I feel unable to concentrate	Mental Fatigue
4. I feel mentally drained	Mental Fatigue
5. I feel alert and focused	Mental Fatigue (negative)
6. I feel worn out physically	Physical Fatigue
7. I feel physically tired	Physical Fatigue
8. I feel somewhat sleepy	Sleep-related Fatigue
9. I feel like closing my eyes and having a nap	Sleep-related Fatigue
10. I feel drowsy	Sleep-related Fatigue
11. I feel wide awake	Sleep-related Fatigue (negative)
12. I feel bored	Boredom
13. I feel detached/uninterested	Boredom
14. I feel uneasy	Negative Affect
15. I feel tense/on edge	Negative Affect
16. I feel irritated and annoyed	Negative Affect

Table 4.11: Preliminary state scale items

4.5.2.2 Factor structure of the state scale

Having accepted the 11 items to represent 3 dimensions of state fatigue and 5 additional items to represent boredom and negative affect, it was considered necessary to investigate the individual item statistics and again subject responses to a process of factor analysis to investigate the clustering of responses into the distinct dimensions. To this end, the 16 state fatigue items were randomly re-organised and put into a questionnaire format consisting of a 5 point Likert scale with verbal anchors ranging from *strongly disagree* (1) to *strongly agree* (5). Twenty psychology students were asked to complete 20 questionnaires each at random time points over a two week period. The sample consisted of 11 females (55%) and 9 males (45%); age range 18 to 23 (mean = 19.84; sd = 1.37). Of the possible 400 questionnaires, 352 were returned, representing a response rate of 88%. These were then treated as individual questionnaires rather than repeated measures data to enable the items to be factor analysed, as the state would be expected to change with respect to current and recent activities.

The data were coded, entered into SPSS and subjected to data reduction. Principal components analysis with varimax rotation was used and eigenvalues greater than one were accepted. A scree plot was examined to help determine the possible number of factors, before rotated solutions were studied.

In the initial solution, 3 factors had eigenvalues greater than one and accounted for 70.3% of the variance. Although most items loaded (>.3) onto all three factors, some psychological meaning could be extracted from the component matrix. The first factor, accounting for 52.4% of the variance, could be described as a general fatigue factor. With the exception of the two physical fatigue items and two of the negative affect items ("I feel uneasy" and "I feel tense/on edge") all items loaded onto this factor (loadings between .31 & .81). The second factor, accounting for 10.1% of the variance, contained the two boredom items, the three negative affect items and one of the mental fatigue items ("I don't feel like making much of an effort"). This factor could be described as a general negative state factor. The third factor, accounting for 7.7% of the variance, had high loadings on the two physical fatigue items (loadings of .80 & .82). However, this factor also contained items from the mental and sleep-related fatigue domains.

Therefore, the three factor solution did not provide a very coherent picture of the structure of the five states. Although, interestingly, the two non-fatigue states did separate out as a distinct factor, providing some support for the distinction between fatigue, boredom and negative affect.

In an attempt to replicate the five domains of the state scale, a five factor solution was forced. This solution accounted for 78.4% of the total variance. The first rotated factor, accounting for 52.4% of the variance, again produced a general fatigue cluster, incorporating mental, physical and sleep-related fatigue items and one of the boredom items ("I feel detached/uninterested"). The second and third factors produced more coherent clusters of boredom and negative affect respectively (accounting for 10.1 & 7.7% of the variance). However, the fourth and fifth factors were less coherent, incorporating items from more than one domain. Therefore, the five factor solution did not support the distinction between the fatigue states, but again differentiated between the fatigue-related items and those generated to represent boredom and negative affect.

Having examined the scree plot, two further solutions of four and six factors were considered. The four factor solution, accounting for a total of 75.2% of the variance, did not offer any coherent structure. However, the six factor solution, accounting for 81% of the total variance, did provide some interesting clusters. The first rotated factor (52.4% of the variance) again contained the majority of the items, with the exception of the negative affect items, providing a general fatigue factor which failed to discriminate between mental fatigue, sleep-related fatigue and boredom. However, the remaining factors provided a more coherent picture. The second, third and fourth factors provided very coherent clusters of the negative affect, physical fatigue and boredom items, respectively, and the fifth and sixth factors principally contained mental fatigue items. This again provided support for the distinction between fatigue, negative affect and boredom, but also provided some limited evidence for differentiating between physical and mental state fatigue. However, none of the above analyses provided any evidence for a distinct factor of sleep-related fatigue.

4.5.2.3 Factor structure of the state fatigue items

To investigate this issue further, the above procedure was repeated using only the 11 state fatigue items to investigate the three factors of state fatigue 'uncontaminated' by other related states. The initial analysis accepted eigenvalues > 1, to investigate an unforced solution, which was followed by a forced three factor solution to determine whether there was any psychometric support for the three intended factors.

In the initial solution, only two factors had eigenvalues greater than one, accounting for 70.3% of the variance. This two factor solution failed to provide a meaningful reduction of the data as the majority of the items loaded onto both factors. Following this, a three factor solution was explored, to investigate the coherence of the three state fatigue domains. Varimax rotation was again used to aid interpretation. The three factor solution accounted for 75.8% of the total variance. Although a number of the items had substantial loadings on more than one factor, the three factors did, on the whole, represent the three dimensions of mental, physical and sleep fatigue (see Table 4.12). The first rotated factor (**sleep**), accounting for 60.5% of the variance, produced a coherent cluster of the 4 sleep-related items (loadings between -.82 & .77). While two of the mental fatigue items loaded on to this factor and factor 3, one of the mental fatigue items, "I feel alert and focused" only loaded onto factor 1 (factor loading -.74). This item was written to reflect the task engagement component of mental fatigue as a central component of mental fatigue is an aversion to further effort. However, despite the aim to write specific items, further consideration of this item suggests that it contains two separate constructs; alertness and focus. It is likely that this item loaded onto the sleep factor because of its reference to alertness. Therefore, this item was removed from further analysis. The second rotated factor (**physical**), accounting for 9.8% of the variance, produced a coherent cluster of the 2 physical fatigue items (both with loadings of .86). This factor, in common with the first, also contained the mental fatigue item "I feel mentally drained" (loading .41). However, this item also loaded onto the third factor of mental fatigue. This third factor (**mental**), accounting for 5.4% of the variance, again produced a clear cluster of the four remaining mental fatigue items (loadings between .47 & .82).

Item	Loading Factor 1: **Sleep**	Loading Factor 2: **Physical**	Loading Factor 3: **Mental**
I feel somewhat sleepy	**.80**		
I feel like closing my eyes and having a nap	**.78**		
I feel drowsy	**.77**		
I feel wide awake	**-.82**		
I feel alert and focused *(removed from further analysis)*	**-.74**		
I feel worn out physically		**.86**	
I feel physically tired		**.86**	
I feel mentally tired	**.59**		**.47**
I don't feel like making much of an effort			**.82**
I feel unable to concentrate			**.70**
I feel mentally drained	**.55**	**.41**	**.47**

Table 4.12: State fatigue items factor loadings (> .4) for forced 3 factor analysis

In comparison with the factor analysis of the trait fatigue scale, the state factors were far less robust. When factor analysed with items representing boredom and negative affect, the fatigue items tended to cluster together, in particular mental and sleep-related fatigue, suggesting a broad factor of state fatigue. However, when analysed independently of the two related states, and forced into a three factor solution, distinct domain related factors did emerge to some degree.

On the basis of the above analysis, an instrument called the 'Feelings Questionnaire" was produced. This was an intentionally ambiguous title which aimed to limit the impact of the negative connotations of the term fatigue. This questionnaire contained 15 items; 4 mental fatigue, 2 physical fatigue, 4 sleep-related fatigue, 2 boredom and 3 negative affect. Descriptive data for each item was calculated from the above sample and is presented in Table 4.13. A copy of the Feelings Questionnaire is available in Appendix 7.

Item / Scale		**Mean**	**Std Dev.**
1. I feel mentally tired		2.62	1.15
2. I don't feel like making much of an effort		2.55	1.15
3. I feel unable to concentrate		2.30	1.05
4. I feel mentally drained		2.49	1.11
Mental Fatigue	**Male**	**2.49**	**0.92**
	Female	**2.48**	**0.98**
1. I feel worn out physically		2.38	1.05
2. I feel physically tired		2.46	1.11
Physical Fatigue	**Male**	**2.40**	**0.97**
	Female	**2.45**	**1.05**
1. I feel somewhat sleepy		2.64	1.17
2. I feel like closing my eyes and having a nap		2.43	1.24
3. I feel drowsy		2.44	1.14
4. I feel wide awake (negative)		3.15	1.14
Sleep Fatigue	**Male**	**2.60**	**1.02**
	Female	**2.55**	**1.07**
1. I feel bored		2.54	1.18
2. I feel detached/uninterested		2.45	1.06
Boredom	**Male**	**2.56**	**1.08**
	Female	**2.34**	**0.97**
1. I feel uneasy		1.93	0.89
2. I feel tense/on edge		1.97	0.97
3. I feel irritated and annoyed		1.88	0.95
Negative Affect	**Male**	**1.94**	**0.82**
	Female	**1.86**	**0.77**

Table 4.13: Descriptive statistics for final scale items and reduced factors (1= strongly disagree / 5 = strongly agree)

4.5.3 Scale Development

Having provided some evidence to further our conceptual understanding of state fatigue, it was necessary to investigate the reliability and validity of the Feelings Questionnaire, before utilising this instrument in the experimental investigation of fatigue. In line with the theoretical framework, changes in state fatigue are an important aspect of fatigue investigation. Therefore, this instrument was developed partially to play a central role in the experiments reported in Chapters Five to Nine.

4.5.3.1 State Scale Reliability

For instruments designed to assess a state, measures of internal consistency are the most sensible estimates of test reliability. Unlike trait instruments, one would not expect there to be high test-retest reliability, as states are not expected to be particularly stable over time. Therefore, it would be inappropriate to attribute a low stability coefficient to inherent limitations of the test, as it may represent an accurate assessment of different levels of any given state on different occasions. Therefore, Cronbach's alpha was calculated for each sub scale as a measure of internal consistency (see Table 4.14).

Feelings Questionnaire Sub Scales	No. of items	Cronbach's alpha
Mental Fatigue	4	0.86
Physical Fatigue	2	0.82
Sleep Fatigue	4	0.90
Boredom	2	0.84
Negative Affect	3	0.81
Whole scale	15	0.93

Table 4.14: Cronbach's alpha coefficient for each state fatigue subscale.

All subscales reached the minimum acceptable level (0.80) recommended by Kline (Kline, 1993). This supports the homogeneity of each subscale and, interestingly, the

'Feelings Questionnaire' as a whole. However, because of the inclusion of the related states of boredom and negative affect, it is not recommended that responses to the entire questionnaire are combined to produce a single index of 'feelings'.

4.5.3.2 State Scale Validation

To provide evidence of the construct and criterion related validity of the 'Feelings Questionnaire' a series of validation studies was undertaken. These included a diary study and a number of experimental manipulations.

Diary Study

Diary Measures: A composite 'questionnaire' entitled Tiredness and Work Demands was produced and presented in an A5 diary format. This contained four parts

- **Part 1 - Criterion related validity:** Brief information about the day's activities, including time starting and finishing work, break time and sleep the previous night.
- **Part 2 - Convergent validity:** This part contained two scales; the current 15 item state fatigue scale and the SAM Fatigue Scale (Pearson & Byars, 1956). The SAM is a unidimensional fatigue scale containing seven items. Each item represents a different degree of fatigue, and a score from 1 to 7 pinpoints the respondent along the fatigue continuum.
- **Part 3 - Divergent validity:** This part contained a modified Positive and Negative Affect Scale: PANAS (Watson & Tellegen, 1985). This contains six items measuring anxiety, depression and fatigue, each of which are represented by 2 bipolar 9 point Likert items. This reduced version was utilised to minimise demand on the respondents (see Hockey, 1996 & Gervais, 2002).

- **Part 4 - Criterion related validity:** This section contained a five item rating scale assessing demands (mental, emotional and physical demands) and opportunities (personal control and personal support).

A copy of a diary page is available in Appendix 8. (These pages were reduced by 50% to produce a week long diary in the form of an A5 booklet.) All participants had previously completed the General Tiredness Scale as part of the trait scale validation study and the data was utilised in the present study to investigate the relationships between state and trait fatigue.

Sample and procedure: Respondents from the trait scale validation study were invited to take part in the diary study for a two week period. This involved daily completion of the diary (either at the end of the work-period or prior to the evening meal if not a work day). The sample consisted of twenty volunteers, of which there were 6 males and 14 females. The age range was 21 to 61 (mean = 39.5; sd = 12.4). With regard to marital status, 12 were married and 8 were single, and half of the participants (10) had children.

Data Treatment: The questionnaire data was reduced into its relevant subscales, as outlined in Table 4.15. The data was then subjected to a range of analyses, within each individual and across the sample.

GTQ (trait measure)	The state fatigue scale	SAM	PANAS
General Fatigue Mental Fatigue Physical Fatigue Morning Tiredness Evening Tiredness Mental Strategies	Mental Fatigue Physical Fatigue Sleep Fatigue Boredom Negative Affect	State Fatigue	Anxiety Depression Fatigue

Table 4.15: Subscale breakdown for validation diary questionnaires

Diary Study Results

Descriptive data: The descriptive data presented in Table 4.16 represent averaged responses to the Feelings Questionnaire (FQ), across individuals, over a two-week period, i.e. the mean subscale scores were calculated for each individual and then the average of these means was calculated for the entire sample. The factor with the highest overall average level is sleep-related fatigue (2.85), which is a little higher than the average level of sleep-related fatigue reported in the first development sample (males mean = 2.60; females mean = 2.55; see Table 4.13). This suggests that the sample in the validation study, composed of working adults as opposed to students, experience a higher level of sleep-related fatigue. Furthermore, while levels of mental and physical fatigue were very similar across the two samples, the working sample in the current validation diary study reported lower levels of negative affect and substantially lower boredom. The finding that working adults (50% of whom have children) report a lower level of boredom than students is not surprising.

Feelings Questionnaire Sub Scales	Mean	Standard deviation
Mental Fatigue	2.49	0.57
Physical Fatigue	2.52	0.74
Sleep Fatigue	2.85	0.54
Boredom	1.92	0.52
Negative Affect	1.79	0.61

Table 4.16: Overall means for the Feelings Questionnaire factors across a two-week period

Table 4.16 also reveals a lower level of standard deviation. However, this is clearly an artefact of the process of averaging mean scores, rather than a reflection of a lower level of variance in state within this sample.

Diary Construct Validity: Having calculated the individual mean values of the five Feelings Questionnaire states, similar values were calculated across the two week period for the single item state fatigue scale (SAM) and the three PANAS scales (anxiety, depression, and fatigue). The correlations presented in Table 4.17 provide some preliminary evidence for the construct validity of the Feelings Questionnaire. The mental, physical and sleep-related fatigue scales of the FQ all significantly correlate (moderately to strongly) with one another, suggesting that those people who are high on one form of state fatigue are typically high on the other types of state fatigue. The moderate to strong correlations between the three state fatigue factors and the SAM provide strong evidence of convergent validity. The particularly strong correlation (r=0.87; $p < 0.05$) between the SAM and FQ sleep-related fatigue was not surprising, as the conceptualisation of fatigue underlying the SAM fits most closely with the sleep-related state fatigue scale, as it emphasises sleep-related descriptors such as wide awake, and extremely tired.

	FQ Mental	FQ Physical	FQ Sleep	FQ Boredom	FQ Negative Affect	SAM State Fatigue	PANAS Anxiety	PANAS Depression
FQ Physical	**63**							
FQ Sleep	**69**	**74**						
FQ Boredom	**54**	07	24					
FQ NA	42	16	10	**70**				
SAM State Fatigue	**63**	**59**	**87**	07	-08			
PANAS Anxiety	08	15	29	22	46	27		
PANAS Depression	-05	-02	22	00	20	30	**91**	
PANAS Fatigue	27	19	**57**	01	03	**63**	**64**	**74**

Table 4.17: Correlation matrix for the Feelings Questionnaire (FQ) factors, SAM state fatigue and PANAS factors (significant relationships in bold; $p < 0.05$).

Regarding the PANAS scale, while Table 4.17 shows a moderate correlation in the expected direction between PANAS anxiety and FQ negative affect, this failed to reach significance ($p = 0.07$). The PANAS depression scale did not significantly correlate with any of the FQ state fatigue factors, which provides evidence of divergent validity. However, following the trait investigation and considering the intuitive relationship between fatigue and depression, a stronger relationship may have been expected. Although some doubt over the choice of the PANAS as a validation instrument may be raised by the very high correlation between the PANAS anxiety and depression scales. Surprisingly, neither the FQ mental nor physical fatigue scales significantly correlated with the single index fatigue score in the PANAS. However, there was a significant correlation between the sleep-related state scale and the PANAS fatigue scale. As with the SAM, the PANAS fatigue scale conceptualises fatigue in terms of 'tiredness' and 'weariness' which is clearly much closer to the conceptualisation of sleep-related fatigue in the current Feelings Questionnaire. This may provide some explanation of the outcome of the various factor analyses of the state fatigue Feelings Questionnaire: None of the solutions for the entire FQ provided a distinct factor of sleep-related fatigue, as the sleep-related items were incorporated either into general fatigue type factors, or loaded onto mental fatigue, boredom etc. This may suggest that the feelings of sleepiness are, in fact, central or common to all types of state fatigue. Although, the factor analyses carried out on the three sleep fatigue factors in isolation did, in fact, produce a distinct sleep factor, which would contradict this. It is currently unclear whether the strong relationships between the FQ sleep-related fatigue and the single index fatigue scores of the SAM and the PANAS are (a) due to these two scales measuring a fundamental fatigue component of sleepiness or (b) simply that the conceptualisation of fatigue within these alternative scales is closer to the distinct domain of sleep-related fatigue than the intended state of 'general fatigue'. This issue will be returned to later in the thesis, during the experimental analysis of the dynamics between demands and fatigue.

The above analysis considered the relationships between these FQ, SAM and PANAS scales on the basis of averaged data over the two week period. This provided information to investigate whether those people who were high in FQ mental fatigue, for example, over the two-week period were also high in fatigue as measured by the SAM, over the same two-week period. However, more fine-grained analyses were required to investigate the relationships between the FQ state fatigue factors and the SAM on each day. This further analysis could address the issue of whether fluctuations in daily state were reflected in both instruments. Table 4.18 presents the correlations between the daily SAM and the daily FQ state fatigue scales. To enable SPSS to carry out these calculations, the analysis was carried out separately for each individual.

This shows that the daily relationships between the SAM and the FQ sleep fatigue scale were consistently strong across individuals. This provides good evidence of convergent validity for the FQ sleep fatigue scale. Interestingly, the relationships between the SAM and the FQ scales differed between individuals. For some participants (e.g. participants 5, 6 and 13) there were strong significant relationships between the SAM and all 3 FQ fatigue scales. Whereas, other participants (e.g. 7, 11, 16) seem to be differentiating between the different types of state fatigue, as they correlate differentially with the SAM. However, from this data it is not possible to determine whether these individual differences reflect a differential ability to distinguish between the different types of state fatigue experience, or whether some individuals experience only a narrow range of demands, which result in mental fatigue (for example) and not physical fatigue while others consistently experience a broad range of demands, resulting in higher levels of mental, physical and sleep-related state fatigue. Although the latter proposition is less likely, an investigation of the criterion-related components of the questionnaire was required to address this question further.

Diary Participant	**Mental**	**Physical**	**Sleep**
Participant 1	**0.89**	0.22	**0.84**
Participant 2	0.41	0.19	0.43
Participant 3	0.21	-0.11	**0.47**
Participant 4	**0.65**	0.31	**0.90**
Participant 5	**0.71**	**-0.82**	**0.81**
Participant 6	**0.87**	**0.71**	**0.78**
Participant 7	0.50	0.46	**0.76**
Participant 8	**0.86**	**0.51**	**0.92**
Participant 9	0.66	0.37	**0.69**
Participant 10	0.50	**0.69**	**0.76**
Participant 11	0.14	0.12	**0.75**
Participant 12	**0.78**	0.50	**0.81**
Participant 13	**0.80**	**0.80**	**0.86**
Participant 14	**0.97**	0.54	**0.90**
Participant 15	**0.78**	-0.43	0.64
Participant 16	0.28	0.38	**0.56**
Participant 17	**0.67**	0.02	**0.84**
Participant 18	**0.61**	-0.41	**0.64**
Participant 19	**0.87**	**0.86**	0.68
Participant 20	**0.80**	0.22	**0.64**
Average correlation	**0.64**	**0.25**	**0.73**

Table 4.18: Correlations between daily SAM and daily FQ state fatigue factors

Diary Criterion Related Validity: To assess the criterion-related support for the subscales, the data were treated in the same way as above: Initially, averaged data across the two-week period was calculated to give an overview of the relationships between the five FQ scales and responses to questions regarding daily demands, control and support and hours of sleep. Following this, a more fine-grained analysis was carried out on the individual data, to consider the relationships between daily experiences and daily states.

Table 4.19 presents the correlations between the averaged states and averaged demands. The significant moderate correlation between the FQ mental fatigue scale and mental demands provides some preliminary criterion-related validity for the FQ mental fatigue scale. This suggests that those people experiencing a high level of mental demands also experience a high level of state mental fatigue. Similarly, the significant moderate correlation between the FQ physical fatigue scale and physical demands provides some criterion-related evidence in support of the physical scale. However, no relationship was found between average sleep fatigue and hours sleep for the two-week period. Although, retrospectively, a more appropriate criterion variable may have been sleep quality, as opposed to number of hours, there was a significant negative relationship between hours of sleep and emotional demands. This may support the use of the 'hours sleep' variable, as one might expect emotional demands to adversely affect sleep.

	FQ Mental	FQ Physical	FQ Sleep	FQ Boredom	FQ Negative Affect	Mental Demands	Emotional Demands	Physical demands	Personal Control	Personal Support
FQ Physical	**63**									
FQ Sleep	**69**	**74**								
FQ Boredom	**54**	07	24							
FQ NA	42	16	10	**70**						
Mental demands	**66**	20	25	33	17					
Emotional demands	**55**	21	17	33	49	**.80**				
Physical demands	37	**59**	35	-15	-06	**46**	41			
Personal Control	**47**	02	21	16	-02	**58**	35	**44**		
Personal Support	**52**	361	23	-09	-05	33	18	40	**63**	
Hours Sleep	-30	-06	-08	**-45**	-16	-40	**-44**	-05	08	16

Table 4.19: Correlation matrix for the Feelings Questionnaire (FQ) factors and the demand and opportunity variables (significant relationships in bold; $p < 0.05$).

Of further interest is the moderate correlation between emotional demands and mental fatigue. This provides support for the proposition, discussed in section 4.4.1.3, that the processes underlying mental and emotional demands are closely linked and, as such, both impact on the development of state mental fatigue.

Finally, the relationships between the FQ scales are worthy of note: With the trait scales, there was a low level of relationship between the subscales, suggesting that these dimensions were fairly independent. Here, however, there are much stronger relationships, a finding which offers further evidence to suggest that the state dimensions are less distinct than the trait dimensions.

The second stage of the criterion-related investigation was to consider the relationships between daily fluctuations in demands and associated changes in FQ states. Therefore, a correlation matrix for each individual was produced (see Appendix 9). With regard to the validation of the FQ mental fatigue scale, twelve of the twenty participants showed positive correlations above $r = 0.3$ between daily mental demands and daily mental fatigue (five correlations were above $r = 0.6$). These correlations reached significance ($p < 0.05$) in nine of these cases, providing some criterion related validity for the FQ mental fatigue scale. Although these relationships were not consistent across all participants, the general trends support the FQ mental fatigue factor, as one would not necessarily expect a very strong relationship between mental demands and mental state fatigue and many other factors may influence state fatigue over the period of a day.

In terms of physical demands and FQ physical fatigue, the relationships were less clear-cut. While there were six participants showing a moderate correlation ($r > 0.3$) between physical demands and physical fatigue (only four of these were above $r = 0.6$ and only these four correlations were significant; $p < 0.05$). Furthermore, three of the twenty participants showed negative relationships between physical demands and physical fatigue ($r = -0.28$, $p > 0.05$; $r = -0.28$, $p > 0.05$; $r = -0.66$, $p > 0.05$). Although none of these negative relationships reached significance, they indicate the complexity of the issues surrounding physical work and physical fatigue. Paradoxically, it does seem intuitive that physically fit individuals may feel invigorated rather than physically fatigued when having completed physical exercise. To investigate the proposition that physical fitness may be a major factor in the relationship between physical demands and physical fatigue, it would be necessary to generate some measure of fitness in participants. However, this information was not collected in the current study.

Having provided considerable construct and criterion-related validity evidence in support of the state fatigue 'Feelings Questionnaire', it remained important to determine whether the FQ was capable of identifying short-term changes in fatigue states. This was not only important with regard to the psychometric evaluation of the construct, but also as the experimental work (reported in the remainder of this thesis) utilised short-term manipulations of mental and physical load to further investigate the fatigue construct. Thus, because of this focus on physical and mental load in the planned experiments, one mental load experiment and one physical load experiment were carried out.

Mental Workload and Fatigue Experiment

Mental workload has been categorised in many different ways. However, it is generally accepted that there are distinctions between the demands associated with time pressure, sustained attention (vigilance) and tasks with a high cognitive load (complex or skills tasks). Furthermore, it is acknowledged that vigilance tasks are the most sensitive to 'time on task' performance decrements and changes in state fatigue (see Desmond & Hancock, 2001). Therefore, a vigilance task would seem an appropriate choice for validating a state fatigue scale. However, because of the differing demands associated with complex or skilled tasks, a complex task would provide an interesting comparison for considering short-term fatigue effects.

Mental Workload Tasks: An adapted version of the Mackworth Clock Test (Mackworth, 1948) was selected to represent a vigilance task. While the original Mackworth task consisted of a broad array of clock type instruments, this task consisted of only one computerised 'clock', represented by 18 small circles forming a large circular structure. Each small circle illuminates in turn in a clockwise direction and, at random intervals, the illumination skips a circle leaving a gap in the continuous pattern. The task of the participants is to monitor the clock and respond by pressing a space-bar to indicate a jump. The programme consisted of a practice session, lasting approximately four minutes, and an experimental session of 30 minutes. During the experimental session, 120 jumps were programmed to occur at random intervals.

To represent a complex task, a numerical reasoning test was selected. This was based upon a fictitious business report which presented participants with a nine page company report, including tables and graphs. The task required participants to extract and manipulate numerical data from the report. Participants were provided with calculators and instructed to complete as many of the questions as they possibly could within the 30 minute period. This task was designed to take longer than 30 minutes, to reduce the likelihood of participants finishing the task early. None of the participants finished before the 30 minute period.

Participants: 34 undergraduate students (15 male and 19 female) volunteered for the experiment. All were between 19 and 28 years of age (mean = 20.4) and had appropriate numeracy and literacy skills.

Design: A repeated measures design with counterbalancing was employed. This required participants to attend 2 experimental sessions and to complete the Feelings Questionnaire immediately before and immediately after each condition.

Results: Having reduced each Feelings Questionnaire into its five factors, pre and post scores were compared for the vigilance task and the complex task. For the vigilance task, *mental fatigue, sleep fatigue, boredom,* and *negative affect* all significantly increased following the 30 minutes of mental work. There was no change in *physical fatigue*. (See Table 4.20.) This provides strong evidence of in support of a multidimensional construct. As one would expect, the biggest increases were in *boredom* and *mental fatigue*, and although there are widely reported effects of prolonged mental work on physical fatigue, changes in physical state fatigue are unlikely over the short time period.

	Pre mental load		**Post mental load**					
	Mean	***SD***	**Mean**	***SD***	**change**	**t**	**Df**	**sig**
FQ Mental Fatigue	2.57	*0.73*	3.39	*0.89*	0.82	5.79	33	$p<0.001$
FQ Physical Fatigue	2.51	*0.94*	2.51	*0.85*	0.00	0.00	33	NS
FQ Sleep Fatigue	2.73	*0.85*	3.49	*0.94*	0.76	3.91	33	$p<0.001$
FQ Boredom	2.64	*0.71*	3.94	*0.89*	1.30	8.57	33	$p<0.001$
FQ Negative Affect	2.30	*0.63*	2.81	*0.88*	0.51	4.63	33	$p<0.001$

Table 4.20: Changes in FQ fatigue scores following vigilance task

For the complex task, a slightly different picture was revealed (see table 4.21). There were again significant increases in *mental fatigue, boredom* and *negative affect*. However, the slight increase in *sleep fatigue* did not prove to be significant in this case and there was again no change in *physical fatigue*. Interestingly, the largest increase was in *negative affect* which, on further investigation, was found to be negatively correlated with performance (number of correct answers) ($r = -0.61$; $p < 0.05$). This suggests that there was a greater increase in negative affect for those people who did less well on the numerical reasoning task. These findings again provide good evidence of criterion related validity for the FQ subscales. Of particular interest is the finding that the vigilance task significantly increased *sleep fatigue*, while the complex task did not. This could be explained by the engaging nature of the mathematical reasoning task, compared with the monotony of the vigilance task and may also highlight the strong subjective similarity between boredom and sleepiness.

	Pre mental load		**Post mental load**					
	Mean	***SD***	**Mean**	***SD***	**Change**	**t**	**Df**	**sig**
Mental Fatigue	2.71	*0.76*	3.14	*0.78*	0.42	2.83	33	*p*<0.01
Physical Fatigue	2.48	*0.86*	2.48	*0.84*	0.00	0.00	33	NS
Sleep Fatigue	2.78	*0.94*	3.02	*0.94*	0.24	1.41	33	NS
Boredom	2.48	*0.74*	3.08	*0.97*	0.60	3.26	33	*p*<0.01
Negative Affect	2.25	*0.52*	2.72	*0.75*	0.46	3.59	33	*p*<0.001

Table 4.21: Changes in FQ fatigue scores following complex task.

Having considered the changes in states for each task, further *t* tests were carried out on the difference scores (post score minus pre score) for the vigilance task compared with a complex task. On the basis of previous research, which has found that vigilance tasks are particularly sensitive to fatigue over short periods, it was predicted that the increases in *mental fatigue, boredom* and *sleep fatigue* would be significantly greater following the vigilance task. The data presented in Table 4.22, does in fact support this hypothesis. This not only provides support for the distinct subscales, but also evidence of the sensitivity of the mental fatigue FQ scale to different levels of mental fatigue.

	Vigilance task		Complex task					
	Mean	***SD***	**Mean**	***SD***	**Difference**	**t**	**df**	**sig**
Mental Fatigue difference	0.81	*0.82*	0.42	*0.87*	0.39	2.21	33	$p<0.05$
Physical Fatigue difference	0.00	*1.03*	0.00	*0.76*	0.00	0.00	33	NS
Sleep Fatigue difference	0.75	*1.12*	0.24	*0.99*	0.51	2.09	33	$p<0.05$
Boredom difference	1.29	*0.88*	0.60	*1.07*	0.69	3.41	33	$p<0.01$
Negative Affect difference	0.51	*0.64*	0.46	*0.75*	0.05	0.35	33	NS

Table 4.22: Difference in changes in FQ scores following vigilance and complex task.

Physical Workload and Fatigue Experiment

While the concept of physical workload is more straightforward than its mental counterpart, physical workload can, nonetheless, be differentiated along similar dimensions. For example, physical endurance, such as running at a sub-maximal rate over a long period, can be compared to vigilance tasks, whereas, physical power work requiring a more demanding load over shorter time period, equates more to complex mental tasks which place a high load on working memory. In the same way that mental fatigue has been shown to be differentially affected by different types of mental workload, physical fatigue may also be differently affected by different types of physical tasks. Although an investigation of the effect of short-term maximal physical load on fatigue would be interesting, the series of experiments, described later, investigate the effects of various sub-maximal physical loads over some time (30-60 minutes). Therefore, it was considered appropriate to validate the FQ physical scale using similar parameters.

Participants: 20 undergraduate students (12 male and 8 female) volunteered for the experiment. All were between 20 and 35 years of age (mean = 22.4; sd = 3.27) and took regular exercise, at least twice a week. A medical questionnaire was used to exclude subjects with a history of hypertension, diabetes, asthma, bronchitis, emphysema, smoking, and any nasal, heart, or lung complaint.

Measures: Participants were required to complete the Feelings Questionnaire immediately before and immediately after the physical load session, as well as the Mental Toughness Questionnaire (Clough, Earle & Sewell, 2002) as used in the development of the trait fatigue scale.

Physical Workload Task: Subjects were required to run on a Powerjog™ treadmill continuously for 45 minutes at 65% of their VO_{2max}[2] which was measured one week prior to the experimental session. This procedure was undertaken to ensure that participants were exercising at equivalent physical loads. This level of workload was selected as it reflects the level of physical work carried out in the later experiments.

Results: Having reduced the Feelings Questionnaires into its five factors, pre and post scores were compared. (See Table 4.23) The only significant change in state was the

---

[2] VO_{2max} fitness testing is an incremental treadmill test to voluntary exhaustion, which provides an indicator of endurance performance potential by measuring cardiovascular function (see Chapter Eight for a description of this process).

highly significant increase in *physical fatigue*. This not only provides criterion-related validity evidence in support of the Feelings Questionnaire, but also provides yet further support for the differentiation between physical fatigue and other types of state fatigue. While the mental load affected all states except *physical fatigue*, the converse was true of the physical load.

	Pre physical load		Post physical load				
	Mean	***SD***	**Mean**	***SD***	**t**	**df**	**Sig**
Mental Fatigue difference	2.22	*0.51*	2.41	*0.35*	1.70	19	NS
Physical Fatigue difference	2.17	*0.76*	3.72	*0.67*	7.68	19	p<0.001
Sleep Fatigue difference	2.61	*0.79*	2.37	*0.59*	1.51	19	NS
Boredom difference	1.75	*0.52*	1.65	*0.51*	0.56	19	NS
Negative Affect difference	1.71	*0.55*	1.65	*0.54*	0.41	19	NS

Table 4.23: State fatigue changes pre and post physical load.

As previously discussed, physical work seems to have an unpredictable effect on fatigue. This is equally true in the context of the current experiment. Although the average change in *physical fatigue* increased, the individual change in *physical fatigue* ranged from -0.5 (less physically fatigued following the experiment) to 3.00 (which represents a considerable increase in physical fatigue on a 5 point rating scale). To investigate the earlier hypothesis about what underlies this individual difference, a Multiple Regression (standard enter method) was carried out including the fitness variable VO_{2max} and the personality variables from the Mental Toughness Questionnaire (see Table 4.24). As predicted, fitness (as measured by VO_{2max}) was found to be a significant predictor of change in physical fatigue. This suggests that the more fit the individual, the smaller the increase in physical fatigue following physical work. This is particularly striking, considering that the physically fitter individuals within the current experiment were required to work at a higher level of absolute physical load to maintain an equivalent level of 'workload'.

	Change: FQ physical fatigue	**MTQ Chall.**	**MTQ Comm.**	**MTQ Control**	**MTQ Confid.**	**B**	**β**	**t**	**Sig.**
MTQ Challenge	.19					.53	.24	1.18	NS
MTQ Commitment	.29	.10				.50	.19	0.88	NS
MTQ Control	.26	.22	**.45**			1.38	.43	1.77	NS
MTQ Confidence	-.16	.00	**.32**	**.45**		-1.15	-.54	-2.37	$p<0.05$
VO2 max	-.31	.31	-.09	.13	-.21	-0.08	-.54	-2.52	$p<0.05$
Mean	1.55	4.15	3.79	3.60	3.66				
Sd	0.90	0.41	0.34	0.28	0.42				R = .69 R^2 = .47 Adj R^2 = .29

Table 4.24: Multiple regression summary table: Predictors of change in physical state fatigue. (Significant correlations in bold; p<0.05)

This suggests that being physically fit does, to some extent, protect an individual from state physical fatigue. Furthermore, *MTQ confidence* was also found to be a significant predictor of change in physical fatigue. As the mental toughness subscales measure aspects of resilience or hardiness, the prediction of change in physical fatigue provides further support for the physical factor.

4.5.3.3 Summary of Feelings Questionnaire (state fatigue) Validation

The diary validity study provided strong evidence of the construct validity of the FQ subscales in its relationships with the SAM state fatigue scale and the PANAS scale over the two-week diary period. More fine grained analysis investigated the daily relationships between the distinct FQ scales and the SAM. This provided further evidence of the construct validity of the FQ scales, particularly the sleep fatigue scale. However, the relationships between the scales were not present for all individuals.

The diary study also provided strong criterion-related support. Averaged over a two-week period, those people reporting a high level of mental demand also experienced a high level of state mental fatigue and a similar relationship was found between physical demands and physical fatigue. However, no relationship was found between average sleep fatigue and hours of sleep for the two-week period. (Although it was argued that hours of sleep may not be the best criterion variable.)

Further evidence of criterion-related validity was generated using short-term manipulations of mental and physical workload. As predicted, there was an increase in mental fatigue following both vigilance and complex mental tasks and the vigilance task resulted in a significantly greater increase in mental fatigue. Similarly, a significant increase in physical fatigue followed 45 minutes of sub-maximal exercise on a treadmill.

As a whole, the evidence collected in the diary study and the experimental manipulations of workload provide both construct and criterion related evidence in support of a multidimensional construct of state fatigue. However, the evidence for a distinct factor of physical fatigue was stronger than the evidence for differentiating between mental and sleep-related state fatigue. Nonetheless, there was sufficient evidence in support of using the FQ in the following experimental investigation.

Part 3:

An Experimental Investigation of the Construct of Fatigue

Chapter 5:

Introduction to Experimental Work

5.1 Summary

Following on from the subjective and psychometric analysis of fatigue, the attention of the thesis now turns to an experimental analysis of the fatigue construct. The theoretical framework of the Compensatory Control model is used to guide the experimental work and interpret the results. As fatigue is hypothesised to occur only under conditions of effortful, goal orientated behaviour the conditions for effective measurement of fatigue involve the evaluation of performance on primary and secondary tasks, subjective state and after-effects. This clearly necessitates a multiple-task environment with a hierarchy of goals. Therefore, the research is conducted using a complex and dynamic process control task, which facilitates the multilevel analysis advocated by the theoretical model. A detailed description of the CAMS simulator task is given.

Four experiments were planned: Experiments 1 and 2 investigate the development of fatigue under different levels of mental workload, Experiment 3 investigates the effects of physical load on fatigue, and Experiment 4 investigated the effects of concurrent mental and physical load on subjective state and after-effects.

5.2 Introduction

Chapters 3 and 4 adopted a subjective and psychometric approach to addressing the conceptual issue of dimensionality of the fatigue construct. Evidence in support of a multidimensional construct was found in existing fatigue questionnaires, a series of interviews and in the psychometric investigation of both trait fatigue and state fatigue, although the evidence in support of a multidimensional construct was much stronger with regard to trait fatigue.

Having carried out a subjective and psychometric analysis of fatigue, and developed workable research tools for measuring fatigue, the attention of the thesis now turns to an experimental analysis of the fatigue construct. The central theme so far has been the multidimensionality of the construct and the extent to which the different types of fatigue can be considered as separate processes. This issue remains a central theme in experimental analysis and the series of experiments culminates in an investigation of the joint effects of mental and physical load on state fatigue. However, prior to this, attention is initially turned towards investigating the dynamics of the fatigue process, focusing on the development of fatigue following engagement with various levels of mental workload. Mental workload, as opposed to physical load or sleep deprivation, was selected as a generator of state fatigue because of the increasing interest in mental demands within various occupations and the impact of different levels of workload. Furthermore, the use of a complex task to manipulate mental workload provides an opportunity to monitor behaviour on a number of levels as recommended by the theoretical model employed.

5.3 Theoretical model

As outlined in Chapter Two, the theoretical model underpinning this thesis is the Compensatory Control model (Hockey, 1993, 1997). This model provides a clear methodological framework for investigating the dynamics of the fatigue process. As previously discussed, an understanding of fatigue has historically been limited by an absence of measurable changes in performance, often in quite extraordinary circumstances. The Compensatory Control model argues that individuals may defy normal expectations of deterioration in performance because of the presence of compensatory control mechanisms, which protect primary task activities. Therefore, any attempt to investigate the nature of fatigue must look beyond simple changes in performance, and consider a broader range of parameters. These parameters, as outlined by Hockey and Meijman (1998), are identified following the predictions of trade-offs between effort and costs, which represent the foundations of the Compensatory Control model.

As fatigue is hypothesised to occur only when an individual engages in effortful goal-directed behaviour, the effects of fatigue should be detectable in secondary tasks, after-effects and subjective state, rather than performance breakdown on the primary task which is unlikely to deteriorate (unless the conditions are sufficiently extreme). Thus, in a high workload environment, an individual has the opportunity to accept or reject the primary goal. If they accept the goal, a *high effort* pattern of decrement is likely: this would involve maintenance of the primary task - at the cost of secondary tasks, performance after-effects and associated changes in subjective state (high mental effort and increased ratings of fatigue). Additionally, while performance may be protected on the primary task, an individual may adjust their primary task management strategy and adopt simpler strategies.

Alternatively, an individual may reject the primary goal, resulting in a *low effort* strategy. In such circumstances, an alternative pattern of performance is predicted, in which performance decrements may be observed on the primary task. However, this low effort strategy is unlikely to generate a state fatigue within the individual and, therefore, there would be little impact on subjective state and after-effects tasks.

The observation of these different patterns of performance breakdown led Hockey and Meijman (1998) to argue that any investigation of the fatigue phenomenon should go beyond the consideration of primary task performance and additionally focus on patterns of latent degradation. Without this broader view, it is argued that it would be impossible to draw meaningful conclusions. Therefore, the current investigation requires a task with multiple goals, options for low effort strategies and opportunities for subjective state measurement.

5.4 Process control: A dynamic, complex task

5.4.1 Single and multiple task paradigms

The use of single, discrete tasks has historically been predominant in workload research. However, the past few decades have seen this practice become increasingly outdated because of limited ecological validity: There is now a wide acceptance that most modern work tasks do not involve the application of single cognitive components, such as working memory or reaction time, but typically require the interaction of different components of cognition (Eysenck, 1984). Therefore, modern workload research has developed dual and multiple task paradigms, which are more representative of the real-world requirement to carry out complex tasks (Wickens & Hollands, 2000).

Ecological validity has been further increased with multilevel tasks, characterised by hierarchical goals. Such tasks typically present participants with a scenario involving a range of different tasks and provide a hierarchy of importance, which (implicitly or explicitly) influences the prioritisation of tasks.

Secondary tasks (those which are not directly related to the primary goal) have an established role in stress and workload research, and have been frequently imposed as a measure of the spare resources which remain untapped by the demands of the primary task. However, despite the prevalence of such tasks in the workload literature, and the growing attention to ecological validity, the secondary tasks employed have typically been simple tasks, with little relevance to real-world hierarchies of complex tasks. For example, rhythmic tapping (Michon, 1966) and random number generation (Baddeley, 1966) are examples of secondary tasks, which of course could be argued to have little real-world relevance (Wickens & Hollands, 2000).

While there are clearly advantages in utilising such tasks, in terms of experimental control and measurement, the theoretical framework adopted here requires a complex multilevel task with a hierarchy of realistic goals. Without this, it would not be possible to identify the patterns of performance predicted by the Compensatory Control model.

5.4.2 Process control simulator tasks

An excellent example of a complex hierarchical task is the dynamic real-world task of process control. This increasingly common task incorporates the three fundamental characteristics of dynamic modern working environments, which have been outlined by Brehmer (1992):

(1) *Time:* This is a critical factor in process control, as such tasks require the operator to continuously update their mental model of the current system, because the problem space is constantly changing, even without any operator intervention.

(2) *Complexity:* Because of the large number of system interactions within a process control task, the state of the system continuously fluctuating and it is not always possible to discern the exact nature of the cause-effect relationships.

(3) *Delays:* The final key characteristic of dynamic tasks is the presence of system delays. These are highly evident in process control environments, as the effects of any intervention will typically be delayed.

Therefore, the adoption of a computerised simulation of a process control task would overcome the ecological validity criticisms of more static or simple tasks. Furthermore, it could also incorporate realistic multiple goals, options for low effort strategies and opportunities for subjective state measurement, which would meet the specific requirements of the current research program and theoretical framework.

5.5 CAMS Simulator: Cabin Air Management System

Having successfully carried out earlier research using a PC-based simulator of a process control task to investigate the effects of occasional nightwork on operator performance and well-being (Sauer, Wastell, Hockey and Earle, 2003) this seemed like an appropriate research tool.

5.5.1 CAMS Background

CAMS was originally developed by Hockey, Wastell and Sauer (1998) as part of a contract with the European Space Agency (ESA) to investigate the influence of various environmental and system parameters on operator performance and well-being. Their aim was to design a simulator tasks which

(1) was truly dynamic, i.e. not operator paced

(2) possessed a high level of face validity in the context of space flight - as it was intended that a high proportion of participants would be astronauts

(3) was highly complex - to maximise ecological validity, and

(4) possessed a number of hierarchical tasks with different priorities - to facilitate an investigation of compensatory control and performance regulation

A full description of the most advanced version of the CAMS simulator (3.0) can be found in Hockey *et al* (1998). This program was utilised for a number of experiments manipulating sleep deprivation, knowledge and training and isolation and confinement (Sauer, 1997). The central feature of this version of the simulator was the operator task of fault diagnosis. However, manipulation of operator workload was a fundamental component of the planned experiments and, as this could be carried out more reliably with an earlier version of CAMS (2.0), this version was adopted. All relevant system characteristics are described below.

5.5.2 CAMS 2.0

The CAMS simulator was modelled on a life-support system for manned space flight. However, the components were simplified so that the findings would be relevant to a broad range of complex working environments, such as energy production, air traffic control, steel making etc. It is based on the control system that regulates the air quality in the self-contained environment of a space cabin. A key feature of the system is that, in common with many process control tasks, the maintenance of its key parameters are highly automated and managed by a number of control devices.

5.5.2.1 The key parameters

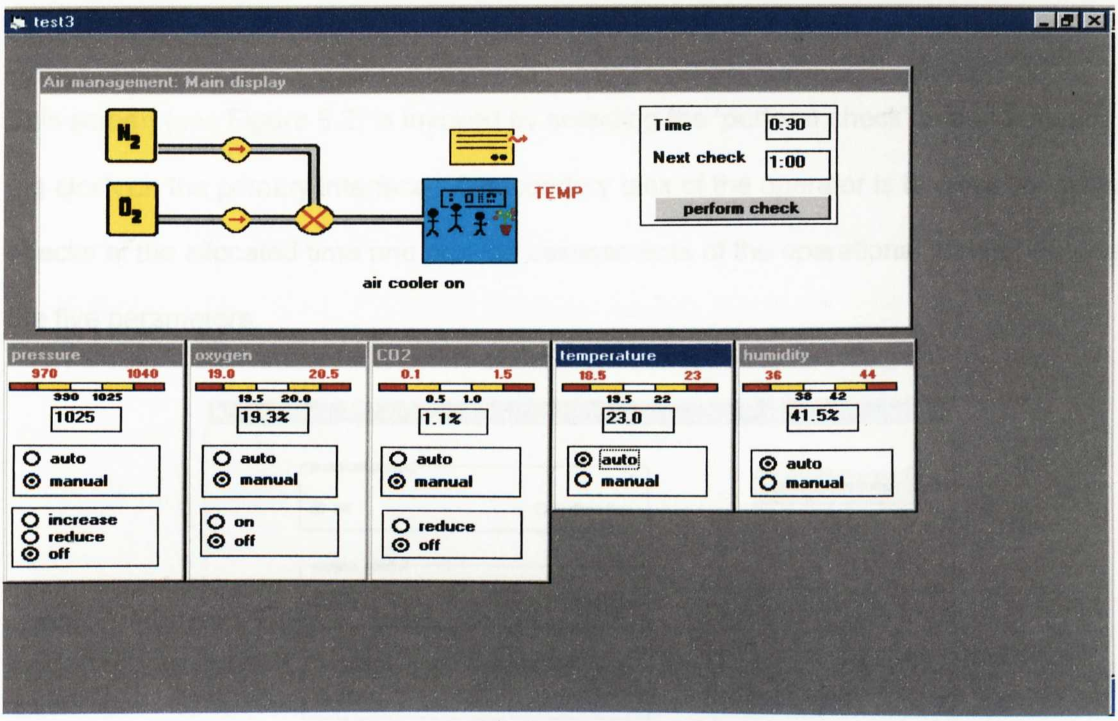
The primary task of the operator is to maintain the quality of the air, firstly by monitoring the level of oxygen, pressure, carbon dioxide, temperature and humidity and, secondly, taking manual control of the appropriate subsystems if any parameter goes beyond the target state. Three operating zones were defined for each of the key parameters. (1) The white zone defines the normal operating range, within which each parameter should, under normal circumstances, oscillate. Each parameter's automatic controllers possess set points at the upper and lower boundaries of the white zones. However, the inherent delays within the system result in occasional drifts into the 'yellow zone'. (2) The yellow zone represents a warning of a potential system failure, particularly if the deviation from the white zone is considerable. (3) The red zone represents a dangerously high or low level of one of the key parameters and a serious risk to the crew and mission.

5.5.2.2 The **CAMS** *operator interface*

The operator interface consists of two screens. The first screen (see Figure 5.1) can be considered to be the primary operator interface. This includes the main display and access to the manual control panels. The second screen is a system status screen discussed below.

The primary operator interface

The main display presents a schematic drawing of the layout of the system. The system consists of a nitrogen tank, an oxygen tank, flow meters within the pipes, a mixer valve, a space cabin and an air-conditioner. Additional system components, not represented in the main display are a dehumidifier and a heater/cooler system. A warning sign is shown when these systems are active, e.g. "air cooler on". In addition to representations of the system components, the main display also incorporates a warning sign for each of the key parameters, which are presented in red when the relevant parameter goes out of range. Finally, the main display contains a panel in the top right-hand corner for the system clock, a prompt for system checks and a button for invoking the system status screen.



Time 0:30
Next check 1:00

pressure
970 1040
990 1025
1025
○ auto
● manual
○ increase
○ reduce
● off

oxygen
19.0 20.5
19.5 20.0
19.3%
○ auto
● manual
○ on
● off

CO2
0.1 1.5
0.5 1.0
1.1%
○ auto
● manual
○ reduce
● off

temperature
18.5 23
19.5 22
23.0
● auto
○ manual

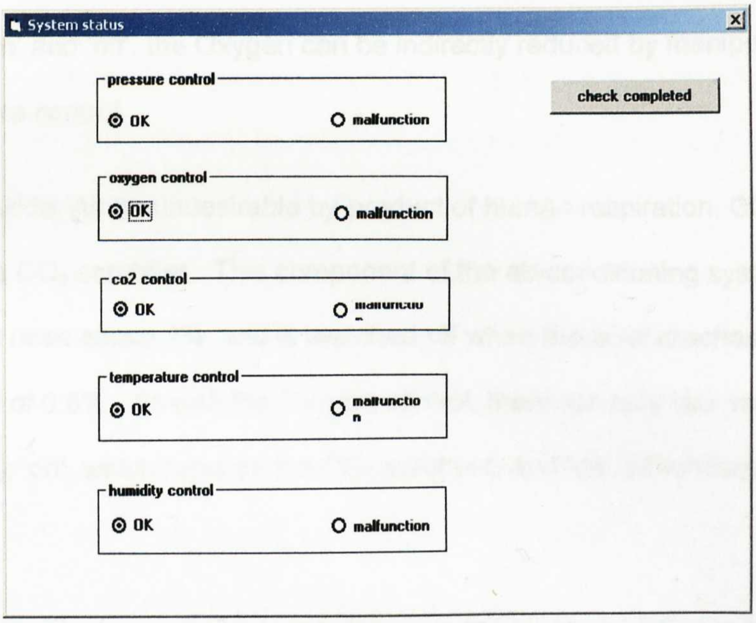
humidity
36 44
38 42
41.5%
● auto
○ manual

Figure 5.1: CAMS primary operator interface

Beneath the main display, there are five control panels (see Figure 5.1). The control panel for each parameter consists of a graphical indicator of the three safety zones and the current levels. It also contains a dialogue box which allows the operator to select the automatic or manual control mode. When manual control has been adopted, the specific control panel will expand to provide options for increasing or reducing the relevant levels. In the case of Oxygen, Carbon dioxide and Humidity, manual options can only be carried out in one direction. For example Oxygen can only be increased. However, because the subsystems are closely coupled, the Oxygen level can be indirectly reduced by accessing the Pressure control panel and selecting reduce. This action will invoke the vent and the levels of both Oxygen and Nitrogen will reduce.

System status screen

This screen (see Figure 5.2) is invoked by selecting the 'perform check' button beneath the clock on the primary interface. A secondary task of the operator is to carry out system checks at the allocated time and provide assessments of the operational state of each of the five parameters.



System status

pressure control: ● OK ○ malfunction

oxygen control: ● OK ○ malfunction

co2 control: ● OK ○ malfunction

temperature control: ● OK ○ malfunction

humidity control: ● OK ○ malfunction

Figure 5.2: CAMS system status screen

5.5.2.3 Basic mechanisms of the life support system

Each of the five parameters is maintained within the target range by using set points, which trigger some regulatory activity once a parameter has reached their specified level. The set points of the normal operating zone were selected to represent the optimum atmospheric environment in a space cabin.

(1) Pressure: The level of pressure is controlled by the operation of the two gas tanks (0_2 and N_2) and the vent function within the air conditioner. When the pressure falls below 990 mbar, the 0_2 and N_2 pumps are automatically switched on, which increases the air

pressure in the cabin. If the pressure rises above 1,025 mbar, the vent is switched on, which reduces the air pressure. Therefore, there are three manual control options, 'increase', 'decrease' and 'off'.

(2) Oxygen: The level of Oxygen is controlled by the Oxygen pump. When the Oxygen concentration falls below 19.5%, the Oxygen supply is turned on. The supply remains on until the concentration of Oxygen reaches 20%. Although the manual control options include only 'on' and 'off', the Oxygen can be indirectly reduced by manipulating the vent via the Pressure control.

(3) Carbon dioxide: As an undesirable by-product of human respiration, CO_2 is regularly reduced by the CO_2 scrubber. This component of the air-conditioning system is turned on when the level rises above 1%, and is switched off when the level reaches the more desirable level of 0.5%. As with the Oxygen control, there are only two manual options for regulating CO_2; 'on', which turns on the CO_2 scrubber, and 'off', which turns off the CO_2 scrubber.

(4) Temperature: The temperature is regulated via the heater and cooler systems. When the cabin air temperature rises above 22°C, the air cooler is switched on, until the temperature falls to 20°C. When the temperature falls to 19.5°C, the heater is switched on, until the temperature is restored to 21°C. Therefore, the manual control panel presents three options, 'heater on', 'cooler on' and 'off'.

(5) Humidity: The level of humidity is regulated by the dehumidifier. This is switched on when humidity rises to 42% and remains in operation until the humidity level is brought down to 38%.

As previously mentioned, one of the characteristics of complex systems is that the components interact with one another. In an attempt to mirror the level of complexity inherent in such systems, and enhance the ecological validity of the results, CAMS was designed to incorporate a number of realistic subsystems interactions (although these were accelerated and magnified because of the artificially short experimental sessions).

5.5.2.4 System failures

Alternative versions of CAMS possess options for a broad range of system failures. These include blockages in the mixer valve, cabin leaks, and leaks in the flow meters. A skilled operator can diagnose these faults by carrying out various relevant system checks. However, within CAMS 2.0, all system failures are caused by a malfunction of the automatic controller. This results in the set points being ignored and, therefore, the corresponding parameters straying beyond the target range. Once an automatic controller has failed, there are no options for maintenance and it will remain in this state until the end of the trial. Manual control must then be adopted, which requires the operator to monitor carefully the relevant level and intervene to maintain the target state. Failures can be programmed to occur within any, or all, of the key parameters, at any time during an experimental session. In the following series of experiments, operator workload is manipulated by the number and timing of system failures.

5.5.2.5 Operator tasks

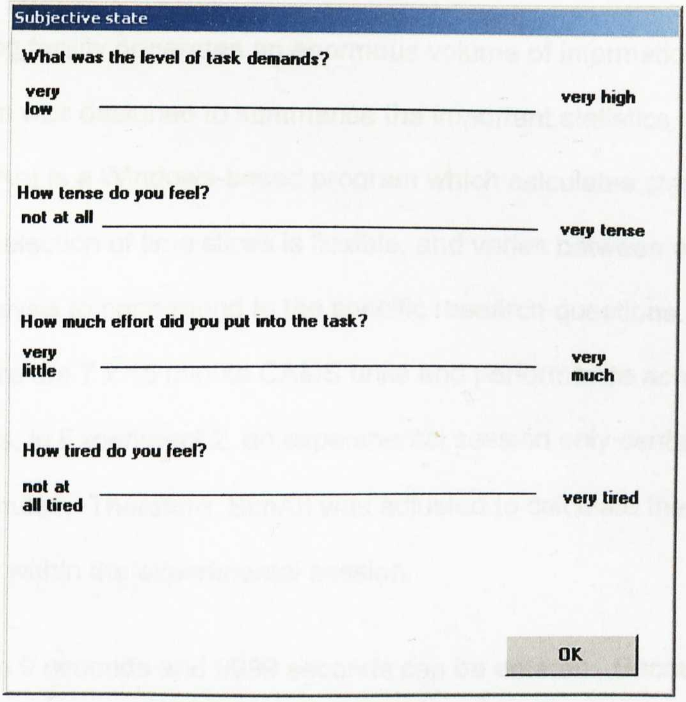
Although the operator instructions varied between experiments, the basic hierarchy of tasks remained broadly the same throughout the experiments. The hierarchy was based on the importance of a task, in relation to crew survival. Although alternative versions of CAMS incorporated a range of primary tasks, such as fault diagnosis, the only primary task in version 2.0 is the *maintenance of the key parameters* within the target state. As stated above, this involves monitoring the five levels and assuming manual control when the operator observes a disturbance in normal system operations.

There are two secondary tasks, *acknowledgement of alarms* and *system status checks*. The acknowledgement of alarms requires the operator to monitor the alarm system and click on any warning signs as quickly as possible. Warning signs can indicate either a false alarm, which vanish upon being clicked, or real alarm, which turns to black when clicked. Real alarms indicate that the relevant system is out of range. False alarms require no further action and genuine alarms require the appropriate intervention. This task provides an indicator of cognitive load by measuring response time. Although this is a relatively simple task, it is widely accepted to be a relevant indicator of primary task load (Wickens, 1992) and possesses some relevance to the real world.

The secondary task of carrying out system status checks is a prospective memory task. This requires the operator to access the system status screen (see Figure 5.2) and adjust any status readings which are incorrect. These reports of the operational state of the system are required at two minute intervals, as dictated by the time prompt, beneath the clock on the main display. This is again a relatively simple secondary task, but with important real-world implications. According to Reason (1990) errors in process control sometimes occur because operators do not remember to carry out actions at specific time points. Performance on this secondary task was scored in two ways. Firstly, mean response delay time and, secondly, according to three categories: correct (+/- 5 seconds), late (+6 to +20 seconds) and omissions (+21 seconds or omissions).

5.5.2.6 Subjective state measures

Within the current theoretical framework, a central component of the study of fatigue is the assessment of subjective state. Two types of subjective state measures are utilised within the current research program. Firstly, CAMS contains subjective state measures embedded within the task. The operator is presented with four visual analogue scales (100mm) and is required to mouse click over each line to indicate their assessment of the *task demands,* their *mental effort* over the previous session and their current level of *anxiety* and *fatigue* (see Figure 5.3). The point at which these measures are taken differs across the experiments to follow. However, they were designed to be unobtrusive and hence were always taken at the end of the experimental session, or at the end of a time period, rather than used as a concurrent workload evaluation. In addition to the embedded state measures, operators also complete a range of pencil and paper questionnaires which vary between experiments and will be described in greater detail in the relevant chapters.



Subjective state

What was the level of task demands?

very low ______ very high

How tense do you feel?

not at all ______ very tense

How much effort did you put into the task?

very little ______ very much

How tired do you feel?

not at all tired ______ very tired

Figure 5.3: CAMS embedded subjective state measures

5.5.2.7 Data logging

The data logging facility provides an entire replication of the experimental session. Data is automatically recorded at 10 second intervals and includes three types of data: system state, operator actions and system actions. (1) Information relating to the state of the system includes the current level of each of the five key parameters (to assess primary task performance) and tank level readings (to assess operator performance on the secondary task of resource consumption). (2) The data relating to operator actions includes every action taken by the operator, such as cancelling the alarms (secondary task) and any manual control options, e.g. manually invoking the dehumidifier. The exact time of occurrence of these actions is also recorded. (3) The data relating to system actions includes all system activity, such as the timing of alarms and automatic regulatory activity, such as turning on the O_2 when the lower O_2 setpoint is reached.

5.5.2.8 CAMS data extraction: Simulator analysis 1.0

As the data logging facility generates an enormous volume of information, a data extraction program was designed to summarise the important statistics. Simulator Analysis 1.0 (SimAn) is a Windows-based program which calculates statistics for selected time slices. The selection of time slices is flexible, and varies between experiments to allow the data analysis to correspond to the specific research questions. For example, in Experiment 1, there are 7 x 15 minute CAMS units and performance across each unit is relevant. Whereas, in Experiment 2, an experimental session only contained one CAMS unit, lasting 90 minutes. Therefore, SimAn was adjusted to calculate the statistics for three time periods within the experimental session.

Any times between 0 seconds and 9999 seconds can be entered. Because of the way CAMS 2.0 logs the information, it is appropriate to have time slices from 0-2000 seconds and 2000-4000 seconds etc. because the software will not double count events such as alarms and system checks. Whereas, the time slices selection of 0-2000 and 2001-4000 will lose information because the second time slices will really start at 2010 seconds. This anomaly occurs because the data file only holds statistics information in 10 second intervals and the parser searching the file starts at a time equal to or larger than the requested time.

This summarised data, plus additional data such as responses to the embedded subjective state measures, can be downloaded into SPSS. Data can either be downloaded from individual files, or groups of data files can be selected and processed as a batch.

5.6 After-effects task

The principal role of the CAMS simulator was, in fact, not to 'measure' fatigue, but to reliably manipulate workload, while monitoring operator state and patterns of performance on the hierarchy of tasks. Following the theoretical framework of the Compensatory Control model, it is suggested that performance decrements on the primary task of CAMS should not be considered as 'fatigue', as they are more likely to represent a low effort strategy. Therefore this simulator is not an appropriate tool to measure actual fatigue.

As a central component of the definition of fatigue is aversion to *further* effort, a thorough evaluation of fatigue should consider performance on after-effects tasks. Therefore, an additional program was required, for studies two to four, which would be presented to participants both prior to and following the completion of the loading tasks.

5.6.1 Fault Finding task

As discussed in the Chapter One, attempts to measure after-effects have met with limited success (Broadbent, 1979; Holding, 1983). The most successful tasks have been those that allow a range of strategies, which differ in the level of processing demands they require. 'Fatigued' individuals have been found to adopt lower effort strategies, even when these are associated with higher levels of risk. Therefore, an after-effects task was sought, which incorporates a number of possible operator strategies, differing in effort required and probability of success.

In an investigation into fatigue and strategy choice, Jongman, Meijman and de Jong (in press) successfully employed a fault diagnosis task, which was an adapted version of the task originally developed by Rouse (1978). They utilised this task to determine the presence and nature of any strategy shifts which occur following performance on a complex scheduling task. Because of the similarities between the scheduling task and the CAMS task, and the successful identification of fatigue after-effects, this was considered to be an appropriate task for the current series of experiments.

The fault diagnosis task is a complex problem-solving task that requires participants to locate a faulty node in a context-independent, logical network (see Jongman *et al* for a full description of their adaptation). This task was again further adapted for the current research program to reduce the required training time and to make the operator interface more user-friendly. A full description of the adapted version is provided below.

5.6.2 Operator interface and task

The operator is presented with a screen consisting of a single network. The network is made up of 30 nodes, linked by a random series of interconnections (see Figure 5.4). The basic premise is that there is a fault in one of the interconnected nodes, and the task of the operator is to identify which of these nodes is faulty. The rules of the network are as follows: Information between nodes is only passed in one direction, from left to right. Therefore, a fault in any single node will only have an impact on those nodes which are connected to it and situated to its right. Contaminated nodes (those which follow on from a fault) will display a red cross when mouse clicked, whereas uncontaminated nodes (those which precede a fault or are not directly connected to it) will display a green tick when mouse clicked.

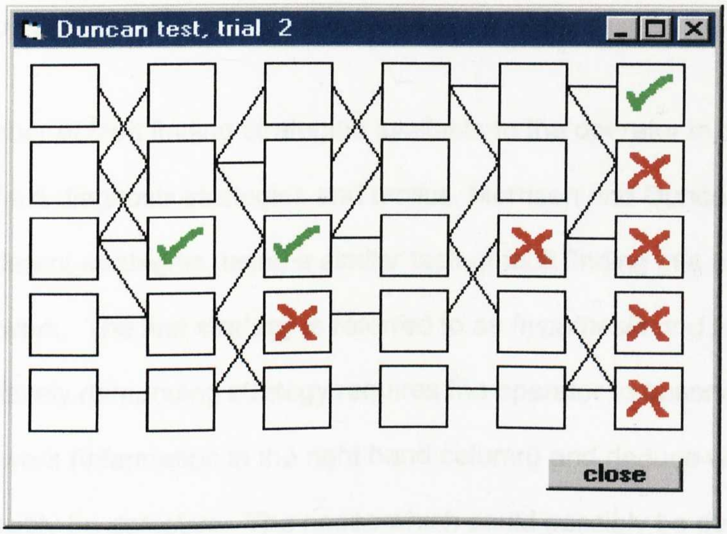


Figure 5.4: Fault finding task operator interface

In finding the fault, operators have two sources of information available to them. Firstly, the column of nodes on the right-hand side of the network represents the network output, which provides information available from the start. A red cross in this column signifies no output from this node (that there is a fault in an interconnecting node which precedes this); whereas a green tick in this column indicates a working output (that no faults are present in any interconnecting node preceding this). The second source of information requires active sampling of the status of a node by clicking upon it. This action will reveal either a red cross or green tick, or may reveal the fault itself, which is signified by a purple 'gremlin'.

As illustrated in Figure 5.4, the operator was initially presented with one working output and four contaminated outputs. The operator has then clicked on four nodes, revealing two contaminated nodes and two working nodes. The information now available reveals that, in this particular network, the fault must be located at either the bottom left node, or the node directly to right of this.

5.6.3 Operator strategies

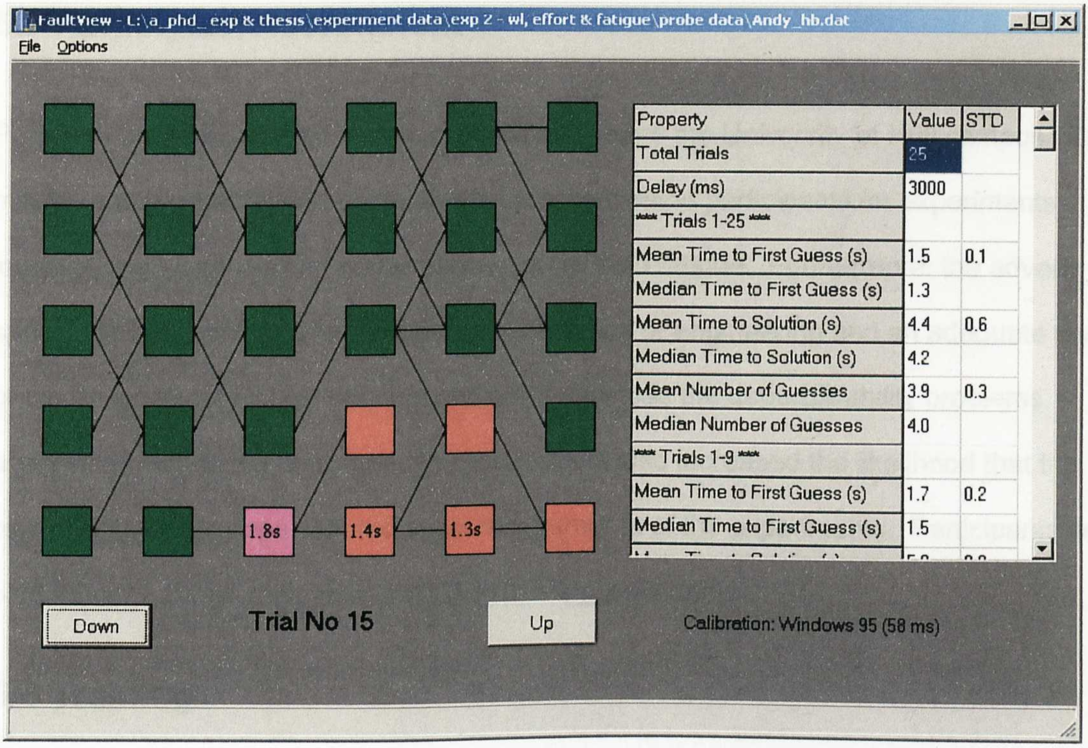
There are a number of fault finding strategies available to the operator in this task. In an investigation of fault diagnosis strategies and tactics, Morrison and Duncan (1988) identified two different strategies, using a similar task of fault finding in a context-independent network. The first strategy is referred to as *hypothesis and test strategy* (HT). This cognitively demanding strategy requires the operator to observe the total output of the network (information in the right hand column) and deduce which of the nodes could possibly be defective. The nodes which could possibly be defective are known as the *FFS: feasible fault set* (Rouse, 1978). Tests are then carried out only on nodes belonging to the FFS. This strategy places a high demand on working memory as participants must keep in mind all those nodes which have been considered. However, this high level of effort is more likely to result in a more efficient diagnosis, as fewer incorrect nodes are likely to be sampled.

The second strategy, again identified by Morrison and Duncan, is known as the *tracing back strategy* (TB). In employing this strategy, operators work back from one of the nodes, that is known to be contaminated, until they find the faulty node. This strategy involves lower effort as it places lower demands on working memory. However, this is typically a less efficient strategy, as more incorrect nodes are likely to be sampled.

In addition to the two strategies identified by Morrison and Duncan, operators also have the option of randomly sampling information. This requires even less effort than the tracing back strategy and is even more likely to result in a high number of errors. Although it is recognised that operators are unlikely to use this strategy throughout their diagnosis, as it is highly inefficient, it is possible that participants may begin with a random node located roughly centrally within the network, and then continue with a tracing back strategy once they have identified a contaminated node. This strategy is more likely to be used when participants are fatigued, because of the low level of effort required. However, unless there are costs associated with this, it is likely that participants may use this lower effort strategy at all times. This would minimise the likelihood of observing any strategy shifts, which is a fundamental component of the current research. For this reason, a time delay was programmed into the fault finding task, so that there was a three second system 'freeze' following each incorrect node selection. This reduced the attractiveness of random guessing as a strategy, making it more likely that this strategy would be employed only when the associated low level of effort became more attractive, i.e. when participants were in a state of fatigue. To determine the optimal time delay, to increase the sensitivity of the task to fatigue states, a small pilot study was carried out. This involved two levels of mental workload, manipulated by the number of system failures occurring during one hour of work, operating the CAMS simulator. A mixed design was used, with repeated measures on the workload variable and independent groups on the delay variable. There were three time delays, 1500 milliseconds, 3000 milliseconds and 4500 milliseconds. Participants were randomly allocated to one of these three conditions and exposed only to this version of the fault finding task during training and their two experimental sessions, which each consisted of one hour of operating CAMS and 25 fault finding networks. Two aspects of fault diagnosis performance were considered. Firstly, average time to first guess, as random strategies were assumed to be quicker, and secondly, number of nodes required to find the fault, as random strategies are likely to be less efficient. Some qualitative data was also obtained, by informally interviewing participants, following their second experimental session. In summary, the 1500 millisecond delay was more likely to engender a random strategy, as the cost of incorrect node selection was not considered to be sufficient to adopt a more effortful strategy. The 4500 millisecond condition was most likely to engender a longer time to the first selection and fewer incorrect selections. The 3000 millisecond condition was the only one of the three delay conditions which differed in the delay to first guess for the high and low workload conditions. Although this suggests that different strategies were used following high and low levels of mental work, this difference failed to reach significance ($p > 0.05$), probably as a result of the small number of participants in this pilot study (n=18; 6 in each condition). However, the qualitative data suggested that the 3000 millisecond delay was likely to be the most sensitive to strategy shifts when fatigued. Therefore, this version of the fault finding task was adopted for Experiments 2 to 4.

5.6.4 Data Extraction

As with the CAMS data logging facility, the data logger for the fault finding task generates a large volume of information. For each of the networks presented, information is recorded regarding the layout of the network, and time and position of each node selected. To facilitate data analysis, an additional extraction program called FaultView was designed. This Windows based program presents two types of information (see Figure 5.5). Firstly, it recreates the specific connections of the network presented for each fault finding problem. As illustrated below, those nodes which would yield a green tick when selected are highlighted in green, the nodes which would yield a red cross when selected are highlighted in red and the pink node indicates the actual fault. Also presented are the nodes selected and the time of their selection, to a 10th of a second.



Property	Value	STD
Total Trials	25	
Delay (ms)	3000	
*** Trials 1-25 ***		
Mean Time to First Guess (s)	1.5	0.1
Median Time to First Guess (s)	1.3	
Mean Time to Solution (s)	4.4	0.6
Median Time to Solution (s)	4.2	
Mean Number of Guesses	3.9	0.3
Median Number of Guesses	4.0	
*** Trials 1-9 ***		
Mean Time to First Guess (s)	1.7	0.2
Median Time to First Guess (s)	1.5	

Trial No 15

Calibration: Windows 95 (58 ms)

Figure 5.5: FaultView Data Analysis Program

Secondly, FaultView also calculates the relevant summary statistics for performance across all 25 networks and across three slices of performance (networks 1-9, 10-17 and 18-25) to allow the investigation of any trends in performance.

Because the reliability of the data logging process in FaultView is paramount, an additional program was developed. *FastCal* is a Windows-based calibration program, which is to be run on any computer used to generate data for the Fault Finding task. This is a simple program, with a runtime of 20 seconds. The calibration constant which is produced is then entered into FaultView to adjust for the impact of different operating systems. This removed the potential error associated with using different machines to collect the data.

5.7 Participant Recruitment

Participants were recruited via adverts placed around the University of Hull campus. To minimise any expectation effects, the adverts requested participants for experiments investigating workload and performance, rather than fatigue. Furthermore, the adverts specified participants with a background in science or engineering and an adequate level of computer literacy. This not only served to minimise the generalisability problems associating with using undergraduate students, it also increased the likelihood that the participants would cope with technical requirements of the experiments. Participants were paid £5 per hour for their time (including training sessions).

5.8 Training

While Experiment 1 utilised only the CAMS simulator task, Experiment 3 used only the Fault Finding task and experiments 2 & 4 used both programs. Therefore, the training sessions varied depending on the programs that were used for that particular experiment.

5.8.1 CAMS Training

The CAMS simulator was designed to represent a realistic process control task. While the complexity of this system is a key feature of the research, in terms of the validity and generalisability of the findings, this has important implications for the training of participants. Successful operation of the system, particularly the manual management of the key parameters, requires a considerable understanding of the system dynamics. Therefore, participants were given sufficient training to enable them to operate the system to an acceptable standard. Although the exact nature of the training sessions varied between experiments, the general CAMS training principles are outlined below.

In general, training involved two 1.5 hour sessions, in which each participant had access to their own PC. The first session focused on an interactive description of the system components and normal operating parameters and provided an opportunity to take manual control of each of the 5 systems. The second session involved a recap of the primary and secondary tasks and then a 30 minute operator test, in which recruits faced 5 system failures, resulting in the need for manual control of all systems. Finally, participants were required to complete a system knowledge questionnaire (see Appendix 10) which assessed their understanding of the system dynamics. Participants who reached the required standard in both the operator test and system knowledge questionnaire were allowed to progress to the experimental stage. However, those who did not reach the standard were given the option of returning for an additional training session, or being paid for their participation so far. Most participants who did not progress through to the experimental stage were overseas students who were limited by their English language capabilities, rather than technical competence. However, a small number of participants struggled with the principles of a dynamic process control task, even following a third training session. These individuals were not invited to participate further.

5.8.2 Fault Finding Training

The training principles for the Fault Finding task were similar to those described above: The nature of the problem was described, focusing on the rules underlying the network and participants were given hands-on experience of all aspects of the system. The various diagnostic strategies were described in detail and participants were encouraged to practice the various strategies available to them. Participants were given a minimum of 2 x 30 minutes of training and, as above, those who did not reach the standard were given the option of returning for an additional training session, or being paid for their work so far.

5.9 Overview of studies

As previously discussed, the central aim of the current research program was to examine the construct of fatigue in terms of its dimensionality. Addressing the question of whether fatigue is a single process or a number of distinct processes requires a series of studies which carefully address the limitations of previous research. Initially, a thorough investigation of the development of mental and physical fatigue must be carried out, in isolation, before any joint effects can be considered. Therefore, the first two experiments represent a systematic investigation of the development of fatigue in a working environment, the third experiment considers the development of fatigue following physical work and the final experiment considers the joint effects of mental and physical work.

Experiment 1

Previous fatigue research has made indefensible assumptions about task demands, their relationship to subjective workload and their links to resultant states. Typically, participants have been required to carry out a task, and then researchers have labelled any performance effects or changes in subjective state as 'fatigue'. This represents a huge leap, which may invalidate any later conclusions and serves to confuse further the fatigue literature.

Therefore, the principle aim of Experiment 1 was to systematically investigate varying levels of task demands, their relationship to subjective workload and any changes in state. This was a fundamental component of the investigation of the fatigue construct, as it is necessary to be definitive about levels of workload before effects of workload on fatigue can be considered.

Experiment 2

Experiment 1 identified appropriate levels of task demand to reflect high, medium and low workload. Experiment 2 then built on these findings to manipulate workload and effort to investigate their respective effects on subjective state and after-effects. A mixed design was used, with repeated measures on the three levels of workload. Although there are many issues surrounding the difficulty of experimentally manipulating effort, this is a crucial component in the theoretical framework of active state regulation (see Chapter Two). Therefore, effort was manipulated between participants on the basis of the 'cover story'. This combination of independent variables enabled a thorough and multilevel assessment of the dynamic process of fatigue following mental work.

Experiment 3

Having established the effects of mental workload on fatigue in Experiment 2, Experiment 3 aimed to investigate the effects of physical workload on subjective fatigue and after-effects. Physical load was manipulated using a cycle ergometer and three levels of individual workload were calculated following maximal capacity fitness testing. In addition to performance and subjective levels of analysis, heart rate was monitored throughout this experiment. The central issues in this Experiment were whether the changes in state were confined to physical tiredness or were more general and whether the after-effects of physical workload were comparable to the after-effects of mental workload.

Experiment 4

This final experiment represented a culmination of the subjective and experimental work thus far. It addressed the central conceptual issue of the thesis, regarding the nature of the fatigue construct. To investigate the interaction between the different types of fatigue, participants were exposed to mental and physical workload concurrently. This experiment utilised a 2 x 2 repeated measures design, with two levels of physical load and two levels of mental load. To correspond with the theoretical framework of the Compensatory Control model, three levels of analysis were addressed: (1) performance on primary and secondary tasks, (2) subjective state, and (3) physiological state.

Chapter 6:

Task Demands, Workload and Subjective State

6.1 Summary

This chapter aims to address some of the limitations of previous fatigue research which has investigated the development of fatigue following mental work. Previous work has often made indefensible assumptions about task demands, assuming that certain levels of demand correspond with specified levels of subjective workload. These assumptions have then formed the basis for arguments about workload and its relationship to fatigue.

Therefore, Experiment One systematically investigates varying levels of task demands and their relationship to subjective workload, performance and changes in state. It is argued that this preliminary experimental work is a fundamental component of the investigation of fatigue, as it is necessary to be clear about levels of workload, before effects of workload on fatigue can be considered.

The CAMS simulator was used as a mental work loading-task, and five levels of workload were manipulated by varying the number of system failures. Participants were required to manage the failed systems and carry out two secondary tasks. They were then required to assess their subjective workload using an adapted version of the NASA TLX.

The results presented in this chapter were used to evaluate the adapted workload questionnaire and to provide a sound basis on which to design the following experiments.

6.2 Introduction

As fatigue is hypothesised to occur under conditions of stress and high workload, an experimental investigation of the fatigue process necessitates a manipulation of one or both of these variables. Because of the methodological and ethical issues associated with generating any of the various states of stress, it is clear that the following experimental investigation would benefit from the manipulation of various levels of workload to generate fatigue. This chapter aims to address some of the theoretical and methodological issues associated with manipulating mental workload, which require careful consideration before addressing the impact of different levels of workload on fatigue.

Interest in the concept of mental workload has increased considerably during the last three decades (Wickens & Hollands, 2000). By the 1980's it had emerged as a central topic in the field of human performance theory (Gopher & Braune, 1984). The primary concerns have been the optimisation of task performance and the measurement of the workload-level associated with particular tasks. While there has also been some interest in long term worker well-being, this has been of secondary importance; possibly due to the fact that the focus of workload research has been to minimise performance breakdown in high-risk work environments, where the potential consequences of poor work design go far beyond the risks to long-term worker well-being. (Although, here it is recognised that long-term worker well-being also has potential consequences for performance breakdown!)

Irrespective of this, workload research has primarily attempted to address acute concerns such as: the point at which performance will suffer; additional tasks that a worker will be capable of handling; and the accurate measurement of workload (Wickens & Hollands, 2000). These are important questions from both applied and theoretical perspectives. However, while they seem to be fairly straightforward, the underlying conceptual issues are complex. At the heart of the problem is the difficulty associated with understanding the relationship between the demands imposed by the task and the operator's ability to cope with these demands (Gopher & Braune, 1984). Moray (1979) outlined three approaches that have been developed to explore these important questions; (1) analysis of the objective parameters of the task – a task demands approach (2) performance-based measures – which focus on performance breakdown, and (3) the subjective appraisal of the level of load an operator experiences. More recently there have additionally been a number of advances taking a psychophysiological approach. The issues associated with these approaches will be discussed below.

6.2.1 Workload and task demands

The first of these is based on the evaluation of workload from the objective demands of the task. This represents a very limited view of the concept of *workload*. Although generating a broadly accepted definition of mental workload has proved difficult (partly due to the multidimensional nature of the construct) modern definitions have typically emphasised the subjective component of meeting work demands, rather than an objective evaluation of task requirements. Therefore, the assumption that workload is somehow equivalent to work demands is fundamentally flawed, as it allows no consideration of the interaction between the task and individual. This distinction is essentially the same as that identified by Kantowitz (1978) between task complexity and task difficulty: While task complexity can be conceptualised as purely a measure of the demands the task imposes on computational processes, task difficulty is a measure of the processing effort required for task performance, which represents an interaction of the task and individual.

A broad number of context-based and state-based factors have been identified as influencing this interaction. One notable moderator is 'task experience', which is known to have an impact on the level of automatic (versus controlled) processing required to meet the task goals (Shiffrin & Schneider, 1977). This example provides a very clear illustration of the distinction between task demands and workload, because the objective demands of tasks do not typically change as one becomes more familiar with the task. However, the increased level of operator experience will reduce the processing effort required to sustain performance and familiar/automatic tasks will be experienced as a lower level of 'workload'. Therefore, attempting to understand or specify workload simply from an evaluation of task characteristics is somewhat limited. Although an investigation of task characteristics would represent an excellent starting point.

6.2.2 Workload and performance measures

The second approach to developing our understanding of workload has been to monitor performance on various tasks and assume that the demands of the task exceed the available resources at the point of overt decrement. Although this method has advantages over the previous method, as it incorporates the interaction between task and individual, the problems associated with such measures are essentially the same problems associated with measuring fatigue: When an individual has adopted a task goal, they are hypothesised to recruit additional resources to meet the increased demand. Therefore, as performance on the primary task will not be subject to linear impairments, this approach has clear limitations (Wierwille & Eggemeier, 1993). This issue is particularly relevant for the current experimental analysis, as performance protection is a fundamental component of the theoretical framework.

A more successful approach to the estimation of workload from performance measures has been the employment of secondary tasks. This methodology assumes that tasks which are additional to the main task can be used as a reliable measure of spare capacity, an approach which has a long history, supported by a large body of literature (see Wickens & Hollands, 2000). As previously discussed in Chapter 5, a broad range of secondary tasks have been successfully employed as sensitive measures of primary task workload and available resources. Within CAMS two types of secondary task are utilised: reaction time and prospective memory. Performance on these measures forms an important part of the multilevel analysis required within the theoretical framework.

6.2.3 Workload and physiological assessment techniques

In addition to the task characteristics approach and the use of performance measures, physiological research has identified a number of techniques which are argued to have potential for evaluating workload. The underlying assumption of physiological workload assessment techniques is that the increases in effort associated with goal oriented behaviour are accompanied by increased central nervous system activity. This is hypothesised to result in a range of measurable changes in psychophysiological functioning, including changes in heart rate and heart rate variability and fluctuations in brain and eye activity (Wierwille & Eggemeier, 1993). Until recently, the most frequently employed of these was heart rate. However, while heart rate is widely accepted as a measure of physical workload, its utility as a measure of mental workload is questionable (Wierwille & Eggemeier, 1993).

Heart rate variability (HRV) however, has increasingly become accepted as superior to simple heart rate as a measure of *mental* effort (Mulder, 1980). More specifically, variability in the mid-frequency band (0.15-0.50Hz) has been widely demonstrated to decrease with task difficulty across both laboratory and field studies (see Wilson and Eggemeier, 1991). For example, Tattersall and Hockey (1995) investigated the relationship between HRV and workload level in flight engineers and concluded that the findings support the use of heart rate variability as an appropriate physiological index of mental effort. Furthermore, these authors found some dissociation between heart rate data and HRV data: While heart rate was found to be more closely associated with task related anxiety (e.g. during the more stressful, but less cognitively demands activities, such as take off and landing), HRV was more closely associated with the mental investment required to carry out complex problem-solving activity. This finding concurs with the earlier conclusion of Wilson and O'Donnell (1988) who, following a review of the literature, concluded that heart rate measures are more accurately utilised as overall indicators of general arousal, than as measures of mental workload. However, although measures of HRV may demonstrate a high level of criterion-related validity as measures of *mental* workload specifically: this technique has been acknowledged as having operational problems associated with the implementation requirements and complex analysis procedures.

In addition to workload measures based on heart rate, a number of measures relating to eye function have been investigated. These include blink rate and blink duration, both of which have a body of evidence to support their application as sensitive measures of visual workload (e.g. Wilson & Fullenkamp, 1991; Hughes, Hassoun, Ward, & Reub, 1990). However, perhaps unsurprisingly, these have been found to be less sensitive to variations in other types of work, such as tasks which emphasise auditory modalities (Casali & Wierwille, 1983).

EEG activity has provided further variables which have been considered as sensitive indices of mental workload. The most commonly employed EEG index is the P300 event related brain potential, the magnitude of which has been found to decrease when task difficulty increases (Kramer, Wickens, & Donchin, 1983).

Therefore, the physiological approach to assessing subjective mental workload has met with great success and, as suggested by Wierwille & Eggemeier (1993) physiological assessment techniques may be most useful when unobtrusive and continuous measures are desirable, particularly where the monitoring of minute to minute variation is required.

However, while physiological techniques have been demonstrated to have considerable utility in an experimental setting, they may prove less useful in real-world settings, were the control of important extraneous variables is difficult.

6.2.4 Workload and subjective assessment

Because of the conceptual and practical limitations with the approaches stated above, the application of subjective workload measures has become widely utilised. The intuitive value of asking an individual about the level of workload they are experiencing cannot be overestimated. So much so, that a number of researchers have argued the point that "If the person tells you that he is loaded and effortful, he is loaded and effortful whatever the behavioural and performance measures may show" (Moray, 1979, p.105). While there are some limitations to a purely subjective approach, subjective workload assessment techniques have been consistently demonstrated to possess an appropriate level of sensitivity to a variety of levels of demand within a range of different tasks (Wierwille & Eggemeier, 1993). Furthermore, their validity has being repeatedly confirmed, as illustrated by Gopher and Braune (1984) who found a strong correlation between subjective measures and the index of task difficulty based on an evaluation of the characteristics of the task.

However, while the practice of subjective workload assessment is widely accepted, there has been some considerable disagreement regarding the dimensions which should be selected for rating. This has led to the development of a number of different subjective measures which are based on different conceptualisations of mental workload. According to Hendy, Hamilton and Landry (1993) the two most notable and widely utilised measures are the Subjective Workload Assessment Techniques and the NASA Task Load Index.

The Subjective Workload Assessment Techniques (SWAT), developed by Reid and Nygren (1988) is, like the TLX, based on a multidimensional model of workload. The three dimensions assessed within this measure are *time load*, *mental effort*, and *psychological stress*. Following a complex card sorting and rating procedure, a single global rating is produced on a scale of 1-100. Although this technique is widely used, it has been generally criticised for the complexity of the time-consuming card sorting procedure and for its reduced sensitivity for low mental workload (Luximon & Goonetilleke, 2001). In response to these criticisms, Luximon and Goonetilleke went on to develop a simplified SWAT and argued that the conventional procedure could be replaced by a simple unweighted average.

In addition to these criticisms, the SWAT has also been criticised on a more fundamental level, for the limited number of dimensions assessed. While this may have been born out of practical necessity, as the card sorting procedure would have become unmanageable with more dimensions, three dimensions may not be sufficient to assess this multidimensional construct fully.

The NASA Task Load Index (TLX), developed by Hart and Staveland (1988) has not been subject to this criticism, as it generates an overall rating of workload from a broader range of six dimensions: *mental demand*, *physical demand*, *temporal demand*, *performance*, *effort* and *frustration*. Ratings on each of six bipolar scales (0-100) are then subject to a weighting procedure, based on a process of paired comparisons (carried out prior to the assessment) which is used to combine the six individual scales into a global score (again on a scale of 0-100).

Although the TLX meets the criticism of limited breadth of dimensions, it has been subject to other criticisms. Nygren (1991) raised serious concerns about the process of paired comparisons, arguing that this scaling procedure lacks an acceptable psychometric basis. However, he did recognise that this process is simple and pragmatic, in comparison to the SWAT weighting procedure.

Although the SWAT and TLX have been dominant, there are also a number of other scales which differ in their dimensionality, and practical application. For example, the Modified Cooper-Harper Scale (Wierwille & Casali, 1983) which is adapted from the first subjective workload assessment technique, the Cooper-Harper Scale (Cooper & Harper, 1969). This 1-10 rating scale was originally designed to assess aircraft handling qualities, although the later version was adapted for more general cognitive tasks. It uses a decision tree to assist the rater in assigning an appropriate value, on a unidimensional rating scale, with the aim of providing a global rating of workload. Yet more simple measures have also been 'developed', such as the Overall Workload Scale (Vidulich & Tsang, 1987). This measure involves a single unidimensional scale of 0-100, zero representing very low workload and 100 representing very high workload.

6.2.4.1 Comparison of subjective measures

There are clearly a wide number of available subjective workload measures. However, selecting between them for various applications raises important questions relating, in particular, to their sensitivity and diagnosticity. While subjective measures in general have been accepted as sensitive to variations in workload level, the relative sensitivity of univariate and multivariate methods remains unclear. In a three-year project to establish guidance for the assessment of operator workload of US army systems, Hill, Iavecchia, Byer, Bittner, Zaklad, & Christ (1992) compared the four measures described above. They found the TLX and the Overall Workload Scale to be superior in sensitivity and to have the strongest operator acceptance. They concluded that, being one dimensional, the Overall Workload Scale may be a useful screening tool for identifying potential 'chokepoints' of workload, but the TLX most useful for obtaining more detailed and diagnostic data.

However, the issue of dimensionality was addressed in much greater detail by Hendy, Hamilton and Landry (1993) who considered whether scaled values from multidimensional scales do in fact provide good estimates of overall processing demands and, furthermore, whether they provide better estimates than judgements made on univariate scales. They concluded that, for an estimate of the *overall* demand on human information processing resources, a univariate workload scale is considered to be more sensitive to varying levels of demand than estimates derived from ratings on a number of different workload dimensions. Furthermore, 'if a univariate workload rating is not suitable, a simple unweighted additive model provides an adequate method for combining the individual factor ratings into an estimate of overall workload' (p579).

An explanation for this counterintuitive, but increasingly accepted, position was provided by Gopher and Braune (1984), who carried out a comprehensive investigation of subjective workload assessments across a broad range of tasks. They argued that, while the mental resources available may be multidimensional, the 'conscious' apparatus appears to follow a single channel. (See Gopher & Braune for a full discussion of these issues).

In addition to the comparison of univariate and multivariate measures, another issue of interest is the frequency and timing of assessments. One of the main advantages of subjective measures is their lack of intrusion into the primary task, as assessments can be obtained following the task session. However, Rehman, Stein & Rosenberg (1983) outlined the limitations of this approach in work environments in which the load is fluctuating, and argued that there is a bias towards the highest point of workload, i.e. that workers may be inclined to forget the easy bits. To address such limitations, Tattersall and Foord (1996) developed the Instantaneous Self Assessment technique (ISA) which presents operators with a short five point rating scale, to which they are required to respond, either physically or verbally, at regular intervals. However, in the context of their particular experimental task, some interference was found on the tracking task both during and following the response, which indicates some competition for attentional resources. This finding does raise some concerns and limits the utility of this measure in certain safety-critical environments. However, the ISA was found to correlate significantly with all three SWAT dimensions, HRV and performance. This clearly provides support for the use of such a measure, particularly in circumstances where the impact of *fluctuations* in workload is an important issue.

Finally, although the reliability of 'post mission' measures could be questionable, due to the time delay between activity and its assessment, Corwin, Sandry-Garza, Biferno & Boucek (1989) successfully utilised post mission subjective ratings, and validated them against videotapes of the mission. They concluded that, although on a long mission the reliability of the subjective ratings may be compromised, delays of up to 15 minutes are not to be considered critical.

In summary, subjective measures in general are widely accepted as reliable and valid techniques for assessing mental workload. In addition to this, subjective measures are practical, minimally intrusive, not disruptive of the primary task, cheap and have a high level of acceptance (Rehman *et al*, 1983). In terms of the distinction between specific subjective measures, univariate measures may currently be held as superior for the purpose of estimating overall workload. However, when the diagnosticity of multivariate measures is advantageous, a simple unweighted average of the scales has been argued to be equivalent to complex weighting procedures.

6.2.5 Experiment rationale

As discussed above, individual operator workload cannot simply be evaluated in terms of the objective characteristics or demands of a task. However, the reliable manipulation of different levels of workload is an extremely important, and frequently overlooked, component of investigating the fatigue process. In the series of planned experiments which constitute the remainder of this thesis, the manipulation of two and three levels of mental workload is fundamental to the interpretation of the pattern of results. Therefore, the experiment described in this chapter aims to investigate the relationship between a broad range of task demands (manipulated via fault combinations in the CAMS simulator), performance on primary and secondary tasks and the impact on subjective state as measured by a workload questionnaire. The primary aim of this experiment is to identify levels of task demands which can accurately be described as representing low, medium and high workload.

6.3 Method

6.3.1 Design

A mixed design was used with two independent variables: (1) Overall session demand and (2) Work-unit demand. Overall session demand was varied at 2 levels; 'high' or 'lower', depending on the total number of system failures occurring within the experimental session. This IV was manipulated between subjects. Work-unit demand was varied at 5 levels (depending on the number of system failures occurring within the specific work-unit). This IV was manipulated within subjects, within the single experimental session.

Although there were only five faults and, therefore, five levels of work-unit demand (ranging from dealing with one fault to five faults) the overall experimental sessions were composed of 7 work-units, each lasting 15 minutes. All participants experienced five levels of work unit demand, but the participants randomly allocated to the low overall session demand condition experienced an additional two work units composed of one and two faults, whereas those allocated to the high overall session demand conditions experienced an additional two work-units composed of four and five faults. Therefore, participants in the low overall session demand condition experienced eighteen faults over the course of the experimental session, while the participants in the high demand condition experienced twenty-four faults.

There were a number of additional considerations in designing the experimental schedule. Firstly, because of the real-world nature of CAMS, faults within the five possible key parameters were not equivalent. They were not only considered to be differentially important in maintaining the safety of the environment, each parameter also oscillated at differing speeds. This clearly impacts on the objective level of demand associated with keeping each fault within range. Thus, to investigate the objective demands associated with each fault, the number of control actions required to maintain the system within range, over a 10 minute period, were calculated. The most outstanding fault was that occurring within the pressure system. Within a 10 minute period, all other faults required between six and eight control actions, whereas the pressure system required thirteen control actions. Therefore, it was determined that participants should be exposed to an equivalent number of pressure faults throughout the session, to reduce the variability in task demand. Of the four remaining faults, oxygen and CO_2 were categorised as *important* faults and control of temperature and humidity were categorised as *less important* and *less demanding*. These categories were taken into account when designing the experimental schedules to ensure that any occurring combination of three faults, for example, would be equivalent, e.g. a pressure fault would not be paired with the two other important and demanding faults of oxygen and CO_2.

Finally, to ensure there was nothing unique about the fault combinations, three different but 'equivalent' experimental schedules were designed for the low overall session demand condition and three for the high overall session demand condition. Consequently, participants were randomly allocated to one of six experimental schedules; three of which were designed to represent low overall session demand and three of which were designed to represent a high overall session demand. (See Appendix 8: Experimental schedules.)

6.3.2 Participants

All participants responded to adverts placed around the Hull University campus. The criteria for selection were (1) to be computer literate, (2) to be studying for, or have completed, studies in a pure or applied science, and (3) to have adequate English-language skills. The advert also stated that some knowledge of process control systems would be an advantage.

The twenty-nine individuals who responded to the advert were provided with more information about the study and invited, via e-mail, to a training session. Twenty-four individuals attended both training sessions, and 22 of these successfully completed the training and took part in the experiment.

Of the 22 participants, 15 were male and 7 were female. Their ages ranged from 20 to 29 with a mean of 24.5 years (sd = 2.2).

6.3.3 Training

Training was given to participants in groups of between 3 and 5 individuals. Because it was considered to be vital that each participant had access to their own PC during the training session, no more than 5 participants could be trained at any one time. Each participant attended a minimum of two training sessions, resulting in a typical training time of three hours.

In the first training session, participants were provided with a verbal explanation of the essential features and functions of the CAMS simulator. They were then given a 'cover story' explaining the nature of the mission and an overview of the primary and secondary operator tasks. They were encouraged to take manual control of each of the key parameters, clicking on each manual option and monitoring the effects on the system.

The aim of this session was to improve their understanding of the system components and the complex inter-relationships between the systems. At the end of this session, the system provided each individual with 'performance' feedback, relating to the amount of time each of the key parameters spent in the red and yellow zones and their resource consumption.

The second session involved a brief recap of the hierarchy of operator tasks and an explanation of the system failures. Following this, participants were presented with a mock experimental session of thirty-five minutes, during which they were faced with 5 system failures, occurring at various intervals. This resulted in the need for manual control of all systems. Performance on this task and on the system knowledge questionnaire (see Appendix 7) was assessed before participants were allowed to progress to the experimental stage.

6.3.4 Independent Variables

This study involved the manipulation of two closely related independent variables.

Overall session demand was varied at two levels. (1) Low overall session demand involved exposure to eighteen faults across the seven work-units. (2) High overall session demand involved exposure to twenty-four faults across the seven work-units. However, the lower demand condition was not expected to reflect true low workload, as there were periods of high demand within the low demand schedules.

Work-unit demand was varied at five levels. (1) Level one presented the participants with one system fault which occurred at three minutes into the 15 minute period. (2) Level two presented participants with two system faults, the first of which occurred at three minutes and the second occurred at nine minutes. (3) Level three presented participants with three system faults, occurring at three, six, and nine minutes. (4) Level four presented participants with faults at three, five, seven and nine minutes and (5) Level five presented faults at three minutes, four and a half minutes, six minutes, seven and a half minutes and nine minutes (see Table 6.1). Therefore, within each 15 minute period, no faults occurred before three minutes and all faults occurring presented themselves before nine minutes.

N°. System Failures	Failure Timings (onset of Fault – in minutes)				
	1st	2nd	3rd	4th	5th
1	3	-	-	-	-
2	3	9	-	-	-
3	3	6	9	-	-
4	3	5	7	9	
5	3	4.5	6	7.5	9

Table 6.1: Timing of onset of system failures for 5 levels of work-unit demand

6.3.5 Dependent Variables

6.3.5.1 Performance related dependent variables

As detailed in Chapter five, the CAMS environment provides a broad range of dependent variables. As the aim of this experiment was to identify levels of task demands which could accurately be described as representing different workload levels, performance on the primary and secondary tasks is of interest.

Primary task performance: This refers to parameter control failures (PCFs), which are defined as the percentage of intervals that each of the five key parameters spends beyond the target state. This is calculated for both errors in the yellow zone (minor errors) and errors in the red zone (major errors). In the current experiment, each work unit of fifteen minutes duration is composed of 4500 samples analysis of the system. For clarity, PCFs are defined as the total percentage of intervals that the five subsystems were out of range.

However, primary task performance is of less interest here than it will be later in the thesis, as errors are so closely related to system failures. In later experiments, the interaction of this variable with other IVs will be of greater interest.

Secondary task performance: (1) Alarm acknowledgement is measured as the time delay between alarm presentation and acknowledgement. (2) Although later experiments categorise system checks as either: on-time, late or missed, the current work-units are only of 15 minutes duration, providing insufficient sampling of this task. Therefore, for the current experiment, only mean responses times will be calculated.

6.3.5.2 Subjective workload assessment

The debate outlined above, regarding subjective workload assessment techniques, has provided a range of considerations for selecting an appropriate technique for the current programme of research. Firstly, the main aim of the experiment is to measure the overall level of workload experienced during the work-units, rather than a continuous measure of the fluctuations in workload throughout the work-units. Therefore, a post task questionnaire was considered to be adequate for this purpose.

The second issue relates to the debate surrounding univariate as opposed to multivariate measures. While univariate measures have been argued to provide a more sensitive measure than scalar estimates along individual dimensions (Hendy *et al,* 1993), the diagnostic capacity of multivariate scales was considered to be an advantage in the current research, as the relationship between fatigue and specific types of work demands may become relevant later. Therefore, a multidimensional scale was utilised, which would be analysed using an unweighted additive model to combine the individual factor ratings into an estimate of overall workload.

The final consideration relates to the dimensions which should be incorporated into the measure. Although there are a number of widely used existing scales (e.g. NASA TLX and SWAT), none of these scales adequately represents the full range of dimensions which have been identified as important. The number of dimensions incorporated into validated scales has been limited by complex weighting procedures. Therefore, as the intention here is to produce a composite measure based on a simple unweighted model, the selection of dimensions has fewer limitations.

Having viewed the available literature, the TLX was considered to be the most appropriate measure, as it incorporated the important dimensions of mental demand, physical demand, time pressure, effort, frustration and performance. However, because of the reduced complexity of the procedure, a number of additional dimensions were also incorporated. These were based on the dimensions identified within the job characteristics model put forward by Wall, Jackson, Mullarkey and Parker (1996). These include monitoring and problem solving demand, and timing and method control.

On the basis of these dimensions, a pencil and paper measure was devised, which incorporated 11 bipolar scales with verbal anchors (mental demand was split into two dimensions of attention and interventions, as these were particularly important components of managing the CAMS task). Participants were required to rate each work unit by marking the relevant point along the scale (see Appendix 9). With the exception of timing and method control and performance (items 5 & 6) the remaining dimensions were reduced into a single index of workload by means of an unweighted average.

Workload data was collected in two ways: firstly, after each 15 minute work-unit participants were asked to rate the work-unit they had just completed and secondly, at the end of the experimental session (of seven work-units) participants were required to complete the same questionnaire but relating to their experience of the experimental session as a whole.

Finally, baseline measures were taken after the first three minutes of (fault free) system monitoring, within each work-unit. At these points participants were asked to rate the level of task demands experienced and the level of effort they expended.

6.3.5.3 Subjective state measures

Subjective state was assessed via an embedded questionnaire which comprised four visual analogue scales, presented at the start and end of each of the seven work units. The measure presented at the start of each unit asked participants to rate current level of *tension* and current level of *tiredness.* The measure presented at the end of each unit asked participants to rate *demand and effort,* along with *tension* and *tiredness.* (See Figure 5.3)

6.3.6 Data treatment

As discussed above, each experimental session was composed of seven work-units. Therefore, participants experienced either one and two system failures twice or four and five system failures twice. For the purposes of analysis, when two measures of any variable were available (e.g. two sets of data for 5 system failures), a mean of the two variables was calculated and used in all tables, figures and statistical analyses.

The main method of data analysis was Analysis of Variance. While this statistical method is considered to be robust against violations of assumptions (Howell, 1992) the performance data presented in the results section were typically positively skewed.

Therefore, to normalise this data, all performance variables were log transformed prior to analysis. However, for clarity, the findings presented in tables and figures will represent untransformed data.

In addition to violations of normality, many of the ANOVAs were found to violate the assumption of sphericity. Where Mauchly's test of sphericity was found to be significant ($p < 0.05$) one of two corrections were applied, in accordance with Field (2000): When the correction factor was greater than 0.75, the degrees of freedom were corrected using Huynh-Feldt estimates of sphericity and when the correction factor was smaller than 0.75, the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity.

6.4 Results

6.4.1 Subjective workload and its relationship to task demand

The primary aim of this experiment was to identify levels of task demand which can be accurately described as representing low, medium and high workload. This therefore required an investigation of the relationship between task demand and subjective workload, as measured by the adapted workload questionnaire.

6.4.1.1 Baseline measures

Although order effects were minimised by counterbalancing the order in which the five levels of work were presented, preliminary analyses were carried out on the baseline measures of demands and effort to evaluate the impact of previous work units on subjective workload assessment. A 2 x 7 mixed ANOVA was carried out on the baseline measures taken at three minutes into each of the seven time periods, i.e. before the first failure occurred.

Baseline demands: As Mauchly's test for sphericity was found to be significant (X^2 (20) = 37.90, $p < 0.05$), the degrees of freedom were corrected using Huynh-Feldt estimates of sphericity ($\varepsilon = .78$). The results revealed that the ratings of demand did not significantly change over the experimental session, $F(4.74, 94.73) <1$, $p > 0.05$, $r = .01$, and, although the ratings of demands were higher in the high workload group, at each time point, there was no significant main effect of overall session demand, $F(1, 20) <1$, $p > 0.05$, $r = .04$. There was also no significant interaction, $F(4.74, 94.73) <1$, $p > 0.05$, $r = .01$.

Baseline effort: As Mauchly's test for sphericity was again found to be significant (X^2 (20) = 45.30, $p < 0.05$), the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .50$). The results revealed that, while the ratings of effort did not significantly change over the experimental session, *F*(3.02, 60.47) <1, $p > 0.05$, $r = .03$, the ratings of effort were significantly higher in the high workload group, *F*(1, 20) = 6.38, $p < 0.05$, $r = .24$. This was an unexpected finding. However, as the mean ratings of effort prior to any fault management were so different (low demand = 6.45; high demand = 9.09), and this difference was sustained across the time periods, this finding should not be attributed to overall session workload, and is presumed to be fortuitous. Furthermore, there was no significant interaction, *F*(3.02, 60.47) <1, $p > 0.05$, $r = .03$.

These findings can therefore be taken as support for the stability of the workload measures across the work-units.

6.4.1.2 Task demands and workload

The composite measure of workload was subjected to a 2 x 5 mixed design ANOVA. As illustrated in Figure 6.1, ratings of workload increased, in both conditions, with increasing system failures. As Mauchly's test for sphericity was found to be significant (X^2 (9) = 35.45, $p < 0.05$), the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .48$). As anticipated, there was a significant main effect of system failure, *F*(1.90, 38.15) = 43.96, $p < 0.001$, $r = .68$. Post hoc Bonferroni tests revealed significant differences between all levels of system failure ($p < 0.05$), with the exception of one and two failures (mean difference = 0.94) and four and five failures (mean difference = 0.72). Although ratings of workload were slightly higher in the high workload group, there was no significant main effect of condition *F*(1, 20) = 1.90, $p > 0.05$, $r = .08$. This provides further support for the workload assessment measure, as there was no evidence of significant carry-over of previous work-units. There was also no significant interaction, *F*(1.90, 38.15) <1, $p > 0.05$, $r = .02$.

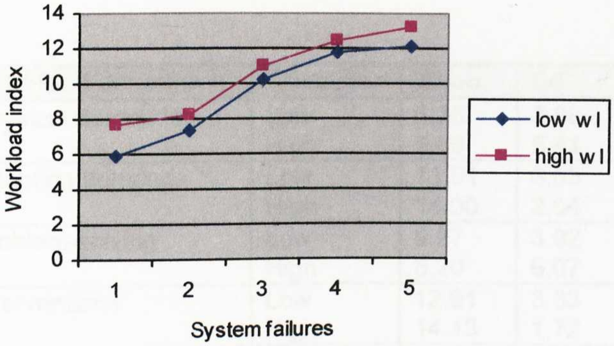


Figure 6.1: Mean composite workload score across number of system failures

In addition to measuring the subjective workload associated with different levels of system failure, the current experiment also aimed to provide some measure of the sensitivity of the workload questionnaire, by comparing workload ratings of the entire experimental session. Although the rating of overall workload for the low workload condition (mean = 10.50, *sd* = 0.77) was lower than the rating of overall workload for the high workload condition (mean = 12.17, sd = 0.64), this difference was found to be not significant, $t(20) = 1.66$, $p > 0.05$.

To investigate this further, a series of independent *t* tests were carried out on the individual workload questionnaire items. With the exception of problem solving, ratings of all individual workload dimensions were higher in the high overall workload condition (see Table 6.2). However, the only one of these differences found to be significant was that of overall session effort, $t(20) = 2.16$, $p < 0.05$.

Workload dimension	Condition	Mean	Sd
Physical demands	Low	6.25	3.88
	High	9.89	5.41
Attention demands	Low	11.91	3.53
	High	14.00	2.04
Problem-solving	Low	9.27	3.92
	High	8.20	6.07
Interventions	Low	12.91	3.83
	High	14.18	1.72
Responsibility	Low	14.18	3.34
	High	16.36	2.15
Effort	Low	12.64	3.23
	High	15.27	2.57
Time pressure	Low	8.55	4.69
	High	11.36	3.58
Frustration	Low	6.36	3.63
	High	7.36	4.61

Table 6.2: Mean individual workload dimensions across two overall workload conditions

6.4.2 Mental demands and changes in subjective state

The data presented in Figure 6.2 reveal different patterns of subjective effects of tension and tiredness across the different levels of system failures. The general picture suggests that, while tension increased with system failures, the level of tiredness appears not to be affected by the number of faults dealt with during the work periods.

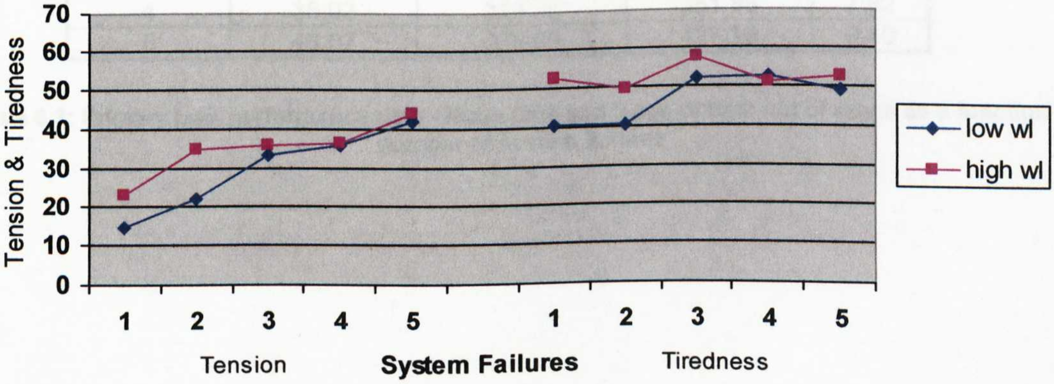


Figure 6.2: Mean tension (left) and tiredness (right) ratings for 5 levels of system failures

In terms of tension, a 2 x 5 mixed ANOVA revealed a significant main effect of workload, $F(4, 80) = 16.35, p < 0.001, r = .45$, no significant effect of condition, $F(1, 20) < 1, p > 0.05, r = .03$, and no significant interaction, $F(4, 80) = 1.40, p > 0.05, r = .06$.

In the 2 x 5 mixed ANOVA, carried out on Tiredness, Mauchly's test of sphericity was found to be significant (X^2 (9) = 22.80, $p < 0.05$). Therefore, the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .64$). There was no significant main effect of system failures on tiredness $F(2.56, 51,22) = 1.95, p > 0.05, r = .08$, no significant main effect of condition, $F(1, 20) < 1, p > 0.05, r = .03$, and no significant interaction, $F(1,20) < 1, p > 0.05, r = .01$.

6.4.3. Primary task performance: Parameter Control Failures

Performance on the primary task was scored in terms of percentage of seconds each parameter was out of range. Mean error scores for the 5 subsystems are presented in Table 6.3. This shows that the number of PCFs increase as a function of system failures.

Number of Failures	Major (red) Errors (secs)	Minor (yellow) Errors (secs)	Total Errors (secs)	%age of time
1	7.64	93.00	100.64	2.24
2	22.48	181.36	203.84	4.53
3	25.32	238.45	263.77	5.86
4	38.02	313.86	351.88	7.82
5	48.07	379.09	427.16	9.49

Table 6.3: Primary task performance data - Mean time and %age of time out of range as a function of number of system failures

Analysis of primary task performance was carried out separately on log transformed major and minor errors. With regard to minor errors, a 2 x 5 mixed design ANOVA revealed a significant main effect of number of system failures, $F(4, 80) = 49.56$, $p < 0.001$, $r = .71$, with an increase in PCFs with increasing system failures.

There was, however, no significant main effect of workload condition, $F(1, 20) = 2.07$, $p > 0.05$, $r = .06$, and no significant interaction, $F(4, 20) < 1$, $p > 0.05$, $r = .06$ (see Figure 6.3).

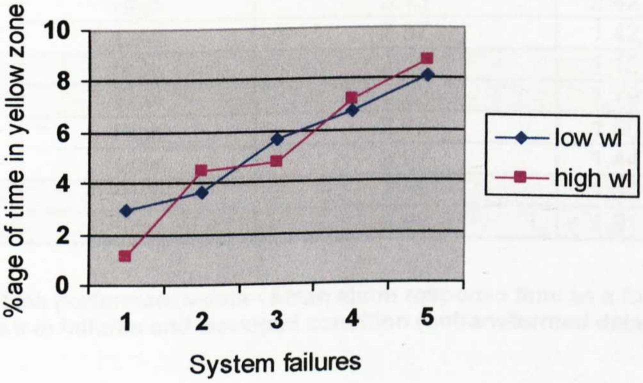


Figure 6.3: Percentage of time in yellow zone (minor errors) across level of system failure and workload condition (untransformed data)

Analysis of the log transformed major errors revealed a similar pattern. There was a significant main effect of number of system failures, $F(4, 80) = 7.21$, $p < 0.001$, $r = .26$, with major PCFs again increasing with increasing system failures, and the effect of overall workload condition on major errors was found to be not significant, $F(1, 20) = 3.62$, $p > 0.05$, $r = .15$. There was also no significant interaction, $F(4,20) = 1.55$, $p > 0.05$, $r = .07$.

6.4.4. Secondary task performance

6.4.4.1 Alarm response time

The first of the secondary tasks was to respond to the system alarms. This data, presented in Table 6.4, was log transformed and subjected to a 2 x 5 mixed ANOVA.

Number of Failures	Workload condition	Mean alarm response time (secs)	Sd
1	High	2.73	1.39
	Low	2.95	1.20
2	High	3.13	2.42
	Low	2.87	1.42
3	High	2.59	1.16
	Low	3.23	1.79
4	High	2.75	1.50
	Low	3.07	1.44
5	High	2.95	1.33
	Low	3.92	1.91

Table 6.4: Secondary task performance data - Mean alarm response time as a function of number of system failures and workload condition (untransformed data)

This unexpectedly revealed no significant main effect of number of system failures, $F(4, 76) = 1.51$, $p > 0.05$, $r = .07$. Furthermore, there was no significant main effect of workload condition, $F(1, 19) <1$, $p > 0.05$, $r = .02$ and no significant interaction, $F(4,76) = 1.31$, $p>0.05$, $r = .07$.

6.4.4.2 System check performance

The second secondary task was to carry out the system checks at the time prescribed by the system clock. This data, presented in Table 6.5, was also log transformed and subjected to a 2 x 5 mixed ANOVA.

Number of Failures	Workload condition	Mean delay to system check (secs)	sd
1	High	4.89	2.24
	Low	2.65	2.08
2	High	4.30	2.73
	Low	3.77	2.48
3	High	4.46	2.06
	Low	4.55	3.61
4	High	4.38	2.94
	Low	2.51	0.85
5	High	5.84	4.40
	Low	4.09	1.85

Table 6.5: Secondary task performance data - Mean delay to system check as a function of number of system failures and workload condition (untransformed data)

As anticipated, this analysis revealed a significant main effect of number of system failures, $F(4, 76) = 2.60$, $p < 0.05$, $r = .12$. However, performance on this secondary task was not as expected, as one would predict an increasing time delay with increasing system failures. Whereas, the differences found were not consistently related to task demand. In terms of overall workload condition, there was no significant main effect, $F(1, 19) < 1$, $p > 0.05$, $r = .00$, and no significant interaction, $F(4, 76) < 1$, $p > 0.05$, $r = .09$.

6.5 Discussion

The overall pattern of results suggests that there is a positive relationship between objective task demands (associated with the various levels of system failure) and subjective workload: As task demands increase, so do subjective assessments of workload, as measured by the composite workload questionnaire. This is an important finding in terms of the current research programme, as the main aim of this experiment was to investigate the relationship between objective task demands and subjective workload and, specifically, to identify levels of task load associated with low, medium and high mental workload. This data was necessary to provide a solid foundation from which to investigate the dynamics of the fatigue process, as the following experiments must be considered to reliably manipulate *different* levels of mental workload.

In terms of the actual level of subjective workload associated with the five different levels of work-unit demand, the mean workload ratings increased with each additional level of system failure from a rating of 7 to 13 on a 20 point scale. However, the differences between one and two system failures and four and five system failures were found to be not significant. Therefore, it can be tentatively concluded that there was little subjective difference between one and two failures and four and five failures. But, in interpreting these results, it is important to recognise that these findings are based on short fifteen-minute work-units. Therefore, the duration of exposure to different levels of task demands may have been insufficient to reflect fully the differences in workload which may result from exposure for long periods. It is in fact most likely that the subjective differences between the different levels of task demands would be greater following more prolonged exposure. Nonetheless, even on the basis of the current data, the subjective experience of managing three system failures *was* significantly different from all other levels.

These results obviously have implications for designing the experimental schedule for the remaining experiments: While three system failures is clearly an appropriate task load to represent 'medium' mental workload, there currently appears to be little to choose between one and two failures, as a representation of low workload, and four and five failures, as a representation of high workload.

To further investigate the issue of what level of task demands might be most appropriate to represent low and high mental workload, the remaining results must be considered in the context of the methodological framework. The Compensatory Control model predicts performance protection on primary tasks, at the expense of decrements on secondary tasks and subjective state. In the current study, however, there was a significant main effect of system failure on primary task performance. But, even with this significant increase in errors, parameter control failures can be considered to be relatively low: Taking account of both major and minor errors, the mean percentage of time spent beyond the recommended parameters ranged from only 2% of time to 9% of time, and a large proportion of this error would be accounted for by the initial drift into the yellow zone, prior to the adoption of manual control. Therefore, although there was a significant effect of system failures on both major and minor errors, primary task performance can, to some extent, be considered to have been protected. However, the secondary task data was not wholly as anticipated: No significant effect of system failure on alarm response time was found, which would have been predicted if there was an increase in workload. However, while there was a significant effect of system failures on system checking, performance decrements did not increase with increasing demand. Therefore, these findings were not considered to be a sufficient basis for making decisions regarding the appropriate level of task demands to represent high and low workload.

Regarding subjective state, while tension was found to increase significantly with system failures, there were no significant effects of system failure on tiredness. Although these findings were mixed, interpretation of this data must take into account the length of work-units: More substantial secondary task effects and changes in subjective state might be expected with more prolonged exposure to work demands, particularly with regard to fatigue as this is argued to develop only as a result of prolonged supervisory-level control. This is unlikely to have been the case in the context of the current short work-units.

In general, these findings support the view that different levels of mental workload can be meaningfully manipulated by varying the number of system failures within the CAMS environment. However, while it was clear that a medium level of mental workload was most appropriately represented by exposure to three system failures, the subjective and performance data suggested that either one or two faults would provide a good representation of low workload, and either four or five faults would provide a good representation of high workload. This issue will be returned to in Chapter 7.

Having addressed the central issue raised by the current experiment, a further consideration was an evaluation of the composite workload questionnaire. Although this instrument was heavily based on existing questionnaires, it was modified to incorporate additional dimensions, and utilised an unweighted average to calculate overall workload. As a result of these modifications, an evaluation of this instrument is appropriate, prior to its adoption for the remainder of the thesis.

A framework for such evaluations was outlined by Wierwille and Eggemeier (1993), who identified a series of dimensions on which to compare workload assessment techniques. This included (1) Sensitivity: the degree to which the measure can distinguish differences in levels of load imposed on operator (2) Intrusion: the extent to which the introduction of the technique causes a change in operator-system performance (3) Diagnosticity: the ability of the measure to discern the type or cause of workload, and (4) Implementation requirements: the need for any specialist equipment, instrumentation, data collection procedures etc.

1. In terms of the sensitivity of the workload questionnaire, composite ratings from the two levels of overall workload were compared. Although there was no significant difference between the 'high' and 'low' overall demand conditions, the mean ratings of workload were in the direction expected. Furthermore, because of the requirement for all participants to experience a full range of work-unit demands, there was in practice only a small difference between the overall loads experienced in the two conditions. The mean difference (of two scale points between the high and low conditions) may in fact represent an accurate reflection of the difference in subjective workload. So, while this difference did fail to reach significance, this finding, coupled with the ratings of work-units (described above) provides some support for the sensitivity of this modified instrument.

2. In terms of intrusion into task performance, this is relatively low with post task questionnaires, such as the current subjective technique. However, this method does also have disadvantages. As discussed earlier in the chapter, post task assessment has been argued to be acceptable when interest is in an overall assessment of workload, rather than in ongoing fluctuations (Corwin *et al*, 1989). Therefore, this approach is particularly relevant in the context of the current

research programme, as workload is simply employed as a fatigue generator (unlike alternative workload investigations, in which the evaluation of workload *throughout* the task is the primary concern).

3. In comparison with other workload assessment techniques, multidimensional subjective instruments are relatively diagnostic, in that they can provide insight into the workload-driving elements within any particular task. The current instrument has been specifically modified to incorporate a wide range of dimensions to maximise its diagnosticity. Therefore, in terms of its evaluation as a workload assessment technique, this instrument provides a major advantage with regard to diagnosticity, which may prove to be an important factor (in later experiments) in furthering our understanding the dynamics of the fatigue process.

4. Finally, paper and pencil questionnaires require very little in the way of instrumentation, and are therefore straightforward and easy to use.

Therefore, this adapted workload questionnaire is considered to be an appropriate tool for the assessment of workload in the following experiments.

Chapter 7:

Mental Workload, Effort and Fatigue

7.1 Summary

The experiment described and discussed in the current chapter aimed to investigate the development of fatigue following different levels of workload and different levels of effort. While the concepts of effort and workload are not always clearly differentiated within the literature, their distinction has important implications for the investigation of fatigue. Therefore, this chapter provides an overview of the effort construct in an attempt to provide some level of distinction.

Following this introduction (and the work presented in Chapter Six) three levels of workload are manipulated by varying the number of system failures in the CAMS task. Two levels of effort are manipulated by varying goal commitment, via different participant instructions.

In summary, this experiment can be considered to have successfully manipulated both mental workload and effort. The pattern of subjective and performance-related after-effects was found to be consistent with the predictions of the Compensatory Control model. This allowed a systematic investigation of the dynamic development of the 'type' of fatigue which develops from effortful engagement with mental work.

7.2 Introduction

A central theme of the current research programme is an investigation of the multidimensionality of the fatigue construct, i.e. the extent to which different types of fatigue can be considered as separate or related processes. Although this remains a central issue throughout the experimental work, attention in Chapters Six and Seven is also turned towards investigating the dynamics of the fatigue process and the development of fatigue following different levels of mental workload. This process requires careful consideration before interactions with physical workload can be investigated. However, the dimensionality issue will also be given some attention here, with regard to the nature of subjective changes in fatigue following mental work.

As discussed in Part One, attempts to advance our understanding of the dynamic processes underlying the development of fatigue have been hampered by the tendency to protect performance. This has resulted in an absence of measurable changes in performance, even in extraordinary circumstances (Hockey, 1993, 1997). The theoretical framework adopted here provides clear guidelines for investigating fatigue, taking a wider view of the regulation of performance and accounting for patterns of latent degradation beyond primary task activities. Highlighted within the model are: decrements in secondary task performance, changes in subjective state and the presence or absence of any after-effects. As previously discussed, this has consequences for experimental methodology, and has led to the adoption of two computer-based tasks, CAMS and the Fault Finding Task. Having systematically investigated the relationship between CAMS task demands and subjective operator workload, the experimental work presented in this chapter builds on these findings to investigate the effects of different levels of mental workload on the various dimensions of subjective fatigue and fatigue after-effects.

Although *workload* was investigated in the previous chapter, the new experiment incorporates an additional factor of mental *effort*. While effort and workload are clearly closely related concepts, their distinction has important implications for the investigation of fatigue. As an illustration of this distinction, consider the position of a CAMS operator: They are required to carry out a series of tasks, each with clear priorities. The objective *demands* of the task can be identified by evaluating the nature of the task requirements, e.g. manual control of one subsystem, which will require eight control actions per minute to remain within the normal operating zone. *Workload*, however, is better conceptualised as a measure of processing requirements, which incorporates an interaction between task demands and the individual and can be influenced by contextual and individual factors. Therefore, in the context of the CAMS task, workload could be described as the nature and extent of processing required by an individual to meet the primary and secondary task requirements. *Effort*, on the other hand, should be considered as a distinct concept from workload, as it is possible that these concepts negatively correlate. Although one would expect ratings of effort and workload to be highly positively correlated, this is dependent on the extent to which the operator adopts the goals inherent within the task. While a motivated operator is likely to monitor performance and invest sufficient *effort* to ensure the task goals are met, an unmotivated operator may recognise that their workload is high, but invest little *effort* as a result of their appraisal of the situation.

The distinction between workload and effort has not always been clear within the workload literature, which may have contributed to some of the anomalies discussed in the previous chapter: As identified by Veltman and Gaillard (1996), the interpretation of workload ratings is clouded by ambiguity as to whether an operator evaluates their workload on the basis of whether they *actually* worked hard, or thinks they *have had* to work hard to perform well.

7.2.1 The concept of mental effort

Mental effort has an indisputable role in the study of human performance and has been identified as one of the most important concepts in the field of information processing and mental workload (Aasman, Mulder & Mulder, 1987). However, the concept of mental effort has not always enjoyed a suitably high profile in the history of psychological research. In common with the concept of fatigue there has been no consistently accepted definition or conception of mental effort, which has allowed it to become a dumping ground for a variety of vague, motivational variables (Fairclough, 1999).

One possible explanation for the lack of interest in the effort concept was put forward by Kahneman (1973) who highlighted the distinction between the study of involuntary and voluntary attention. There was a substantial early interest in *involuntary attention* which focused on aspects of the environment to which humans involuntarily attend (Berlyne, 1960). For example, a loud unexpected noise will typically 'gain our attention', irrespective of our current focus or ongoing mental activity. An important component of this study of involuntary attentional processes is the associated involuntary changes in arousal. This led to an interest in so-called *intensive* aspects of attention. However, despite the early interest in such issues, cognitive psychology shifted away from this line of research emphasising instead the process of voluntary attention, i.e. the direction of attention towards stimuli which are deemed to be relevant to specific tasks. This research was dominated by structural models of information processing systems and was conducted with little or no reference to arousal or to the intensive aspects of attention. Kahneman (1973) however, argued that intensive aspects of attention must be considered in dealing with voluntary as well as involuntary attention. In this seminal work, he distinguished between the concept of effort and the even more elusive concept of arousal, associating arousal with wakefulness and defining effort in terms of what an individual is *doing* at any one time.

Despite this influential work, the concept of mental effort continued to receive little attention. One possible explanation for this was highlighted by Hockey, Gaillard and Coles (1986), who suggested that the concept of mental effort is unfortunately positioned equidistantly between biological and cognitive traditions in psychology: While the biological perspective defines mental activity with respect to psychophysiological changes occurring within the central nervous system, the cognitive perspective has been dominated by the computer metaphor, which emphasises information processing structures. Although these 'wet' and 'dry' approaches each have a role to play in advancing our understanding of mental activity, neither can adequately account for variability in human information processing. Central to this variability is the notion of mental effort, in particular, its active regulation and the strategic control of behaviour. Such issues have been framed within the modern field of cognitive energetics (Hockey *et al*, 1986), which aims to bridge the gap between the biological and cognitive traditions. In fact, the Compensatory Control model (adopted here) is an excellent example of the work carried out from this perspective.

While interest in cognitive energetics has brought the issues of mental energy, mental effort and mental resources into focus, effort continues to be conceptualised in many ways and there remains no single, unified definition of effort. However, there are a number of common elements which can be identified across the majority of definitions, e.g. the volitional control of mental effort, costs associated with investment of effort and the impact of effort investment on task performance. These components all relate to the general concept of effort in relation to attention-demanding controlled information processing (e.g. Kahneman, 1973; Mulder, 1980; Shiffrin & Schneider, 1977).

However, there is an alternative approach to the concept of mental effort, which goes beyond effort as 'executive resource control' and views mental effort as synonymous with state regulation. According to Hockey and Hamilton (1983) several states of the human information processing system can be defined which influence the availability of cognitive resources. When the actual state does not correspond to the optimal state required to perform a task, the mismatch is registered and a number of possibilities are available, e.g. adjustment of the current state or adjustment of the target state. Effort is the central energetical construct involved in state regulation.

These two approaches or conceptualisations of mental effort are, however, not mutually exclusive and can be integrated into a single framework. For example, Mulder (1986) distinguishes between the two purposes of investing effort (1) performance improvement and (2) state effort: protection of performance in the face of physical and environmental stressors. Therefore, effort can be argued to be related both to the amount of control processing required and to the amount of required change in state.

7.2.2 Mental effort and Fatigue

While the mental effort literature is complex and sometimes contradictory, the role of effort in the development of fatigue is much clearer. As discussed in Chapter Two, the Compensatory Control model makes clear predictions with regard to effort and fatigue: fatigue is argued to develop as a consequence of effortful task engagement. Therefore, while workload manipulation plays an important role in generating fatigue, it is the investment of *effort* itself which is central to the current programme of research. As originally highlighted by Kahneman (1973) one can, to some extent, manipulate effort via workload. He proposed that the effort invested in a task is determined mainly by the intrinsic demands of the task, and that voluntary control over effort is quite limited. While this may be true at low levels of workload (consider low level 'data-limited' tasks; Norman and Bowbrow, 1975) the investment of effort at higher levels of workload is clearly (at least partially) discretionary. An illustration of the ability to control effort was highlighted by the work of Frankenhaeser (1978), who identified different patterns of arithmetic performance and costs under noise: One pattern was associated with performance decrements and no change in adrenalin or subjective effort. The other pattern was associated with performance 'protection' and increased levels of adrenalin and subjective effort. The workload of these two 'groups' was the same, and therefore, it could not be argued that the effects of task performance were purely a result of the workload.

Therefore, a full investigation of the fatigue construct needs not only a manipulation of workload, but also a manipulation of effort.

7.2.3 Experiment rationale

As outlined above, effort investment can, to some degree, be influenced by the intrinsic characteristics of the task, for example, the number of system faults which require manual management. However, it is also widely recognised that the allocation of processing resources, or the investment of effort, can be influenced by other factors including (1) effects of arousal, (2) enduring dispositions which reflect the rules of involuntary action (Kahneman, 1973) and (3) goal commitment (Mulder, 1986; Latham & Locke, 1991). For the purposes of the current experiment, the most suitable way of manipulating effort is clearly via the manipulation of goal commitment.

Therefore, the current experiment aims to investigate the development of fatigue following three different levels of workload and under different levels of effort.

7.3 Method

7.3.1 Design

A mixed design was used with two independent variables: (1) Workload and (2) Effort. Workload was varied at three levels: *low*, *medium* and *high*, determined by the number of system failures occurring within the experimental session. This IV was manipulated within subjects, resulting in three experimental sessions. Effort was manipulated between subjects, on two levels: *normal* and *high* effort.

Each experimental session lasted approximately two hours: Participants were required to complete 25 Fault Finding networks before and after 100 minutes of CAMS operation.

To minimise any carryover effects, a minimum delay of three days between experimental sessions was enforced. The maximum delay was seven days.

7.3.2 Participants

Of the 22 participants who successfully completed Experiment 1, twenty volunteered to continue with the series of experiments. The two participants who no longer wished to take part said they could not commit to the demanding requirement to complete additional training and three experimental sessions.

In recognition of the importance of adequate sample sizes, additional participants were recruited. Adverts were again placed around the University of Hull campus requesting participants from a background of pure or applied science, with computer literacy and adequate English-language skills. Fifteen individuals responded to the advert and were invited to a preliminary training session. Twelve of these individuals attended both training sessions, although only 8 successfully completed the training and took part in the experiment.

To balance the experimental experience across the groups, the original 20 participants were randomly allocated to either the 'normal 'or 'high' effort condition and the same random allocation process was carried out for the 8 new participants.

Of the 28 participants, 18 were male and 10 were female. Their ages ranged from 20 to 35 with a mean of 23.8 years (sd = 3.3).

7.3.3 Training

The training principles adopted for Experiment 1 were maintained, i.e. participants were trained in small groups, with sole access to a PC to facilitate learning. Two parallel training schedules were undertaken to account for the differences between the new and experienced participants.

1) *Experienced participants*: Those participants who had previously taken part in Experiment 1 were required to attend two additional training sessions. As there was a time delay of approximately six weeks between Experiments 1 and 2, training session 1 began by reminding participants of the main operating principles of CAMS, and the hierarchy of operator tasks. They were then provided with 20 minutes hands-on system operation, during which their performance was evaluated to ensure minimum operating standards were maintained. All participants reached the acceptable standard for primary task performance. Participants were then introduced to the Fault Finding Task, including a description of the basic features and operating principles. This was followed by an explanation of the possible strategies available to the participants (outlined in Chapter 5). Finally, participants were asked to work through a series of 25 networks, employing all potential

strategies to meet the dual goals of (1) finding the gremlin as quickly as possible, whilst (2) selecting the minimum number of nodes to find the fault. These opposing goals were deliberately not prioritised.

The second training session for the experienced participants required them to complete 25 networks, again following the dual goals of speed and parsimony. Performance on this task was also monitored to ensure full understanding of the task.

2) *New participants:* The new participants were also subjected to two training sessions. These two sessions followed the CAMS training schedule outlined in 6.3.3. However, at the end of each session, participants were trained on the Fault Finding Task, as described above. Therefore, new participants were subjected to the same training as existing participants, although, for this group, each of the two training sessions incorporated both CAMS and Fault Finding training. In all, both groups were exposed to a total of 1.5 hours CAMS training and one hour of Fault Finding training.

7.3.4 Independent Variables

This study involved the manipulation of two IVs: workload and effort.

Workload was varied at three levels; *low*, *medium* and *high*. The number of system failures selected to represent each of these levels of workload was based on the findings of Experiment 1. (1) Two system failures were selected to represent *low* workload: Although there was no difference between one and two failures in terms of performance and subjective evaluation, a number of participants stated that managing a single failure was "uncomfortable" and that it was particularly difficult to remain focused and engaged during these low demand work-units. Therefore, to avoid the possibility of an underload condition, two system failures were programmed to represent low workload. (2) Three system failures represented *medium* workload, and

(3) all five system failures represented *high* workload. Therefore, all participants experienced three randomly ordered experimental sessions, each representing one of the three levels of workload. To ensure there was nothing unique about the selected fault combinations, two equivalent experimental schedules were designed for each level of workload, and participants were again randomly allocated to schedule A or B (see Table 7.1).

	Fault presentation			
Workload condition	*Schedule A*	(mins)	*Schedule B*	(mins)
Low workload	Oxygen	*3*	CO_2	*3*
	Temperature	*50*	Humidity	*50*
Medium workload	CO_2	*3*	Oxygen	*3*
	Humidity	*30*	Temperature	*30*
	Pressure	*50*	Pressure	*50*
High workload	Humidity	*3*	Temperature	*3*
	CO_2	*12*	Oxygen	*12*
	Pressure	*30*	Pressure	*30*
	Temperature	*40*	Humidity	*40*
	Oxygen	*50*	CO_2	*50*

Table 7.1: Experimental schedule: Possible fault combinations

Across all levels of workload, the first fault was presented at three minutes and the final fault was presented at 50 minutes. Therefore, for the second half of the duration of CAMS operation, all faults were present.

Effort was varied at two levels via the 'pre-mission' instructions: (1) Those randomly allocated to the *normal* effort condition were provided with the standard instructions, including a cover story of a mission involving 'critical biological, chemical and medical experiments'. They were also reminded of the hierarchy of primary and secondary operator tasks and that a high level of primary task performance was vital to the success of the mission. (2) The instructions for the *high* effort condition were the same. However, in an attempt to maximise effort by increasing commitment to the task, the participants in the high effort condition were told that the success of the current programme of research depended entirely on them working very hard to achieve the task goals to the best of their ability, and that there would be no point in taking part unless they tried their absolute hardest. Therefore, while the workload and demands of the task was constant across both effort conditions, the high effort group should have been more motivated to maintain a high level of performance.

7.3.5 Dependent Variables

7.3.5.1 Performance related dependent variables

In addition to the primary and secondary task performance variables provided by the CAMS environment (see 6.3.5.1) the current experiment incorporated the Fault Finding task into the experimental session. Both prior to and following CAMS operation, participants were required to complete 25 trials of the Fault finding task. This provided a series of additional dependent variables, which were extracted from the Fault Finding programme and summarised by the FaultView data analysis programme. Summary statistics available were mean and median measures of (1) time until first guess (2) time until solution and (3) number of guesses.

7.3.5.2 Subjective measures

As with Experiment 1, participants were required to complete a pencil and paper workload questionnaire (see 6.3.5.3) at the end of the CAMS task.

Subjective state was assessed in two ways. (1) Visual analogue scales embedded within the CAMS task measured ratings of demand, effort, tension and tiredness. (2) State fatigue was measured via the 15 item Feelings Questionnaire, presented in Chapter 4.

Finally, there were two effort-related measures, employed to investigate the impact of the effort manipulation (see Appendix 10). The first of these was a 'task engagement' questionnaire and required participants to mark along three visual analogue scales, which related to their level of engagement with the primary and secondary tasks. The second was a ratio-based effort questionnaire, which asked participants to divide their effort allocation into three, to represent the way they allocated their effort between the primary and secondary tasks.

7.3.6 Data Treatment

Analysis of Variance was again the main method of data analysis. Continuing with the principles of data treatment established in Chapter 6, all performance variables were log transformed prior to analysis. However, to facilitate interpretation, the findings presented in tables and figures represent untransformed data.

Furthermore, where sphericity has been violated, degrees of freedom were corrected using either Greenhouse-Geisser (< 0.75) or Huynh-Feldt (> 0.75) estimates of sphericity, in accordance with Field (2000).

The questionnaire data was reduced into the relevant factors: The workload questionnaire was reduced into a single index of overall workload; state fatigue was reduced into five factors of mental, physical and sleep-related fatigue, boredom and negative affect, and difference scores were then obtained by subtracting pre- from post-factor scores.

The data from the Fault Finding Task was extracted using the Data Extraction programme and entered into SPSS. As this data was found to be skewed, pre and post variables were log transformed, prior to the calculation of difference variables.

7.4 Results

7.4.1 Workload, effort and task performance

The primary aim of this experiment was to investigate the development and nature of fatigue following three different levels of workload and two different levels of effort. However, prior to analysis of subjective fatigue and after-effects, it was important to consider primary and secondary task performance, and the impact of the effort manipulation.

7.4.1.1 Primary task performance

Performance on the primary task was again scored in terms of percentage of time each parameter was out of range. Parameter control failures were categorised as either major errors (time spent in the red zone) or minor errors (time spent in yellow zone). (See Figures 7.1 & 7.2.)

A 3 x 2 mixed ANOVA was first carried out on the log transformed major errors. As Mauchly's test for sphericity was found to be significant (X^2 (2) = 40.11, $p < 0.05$), the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .55$). As expected, this revealed a significant main effect of workload $F(1.12, 28.90) = 14.26$, $p < 0.001$, $r = .35$, and a significant main effect of effort condition, $F(1, 26) = 5.09$, $p < 0.05$, $r = .16$. Interestingly, this analysis also revealed a significant interaction, $F(1.12, 28.90) = 5.38$, $p < 0.05$, $r = .17$. As illustrated in Figure 7.1, primary task performance can generally be considered to have been protected, with only a small proportion of time spent in the red zone, across all conditions. However parameter control failures did increase with increasing workload and performance was poorer in the normal effort condition. In terms of the interaction between workload and effort, the impact of the high effort condition was much greater under high workload, which suggests that greater effort (or task commitment) can lead to better performance protection in conditions of high workload.

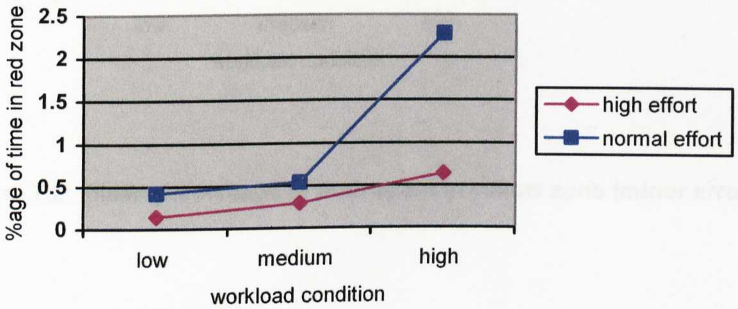


Figure 7.1: Mean percentage of time spent in red zone (major errors)

A similar pattern of results was found for minor parameter control failures, which were subjected to a 3 x 2 mixed design ANOVA. As Mauchly's test for sphericity was again found to be significant (X^2 (2) = 21.44, $p < 0.05$), the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .63$). This analysis again revealed a significant main effect of workload $F(1.26, 32.99) = 41.58$, $p < 0.001$, $r = .61$, with parameter control failures increasing with increasing workload. But, although minor errors were higher in the normal effort condition at each level of workload, this difference did not reach significance, $F(1, 26) = 1.98$, $p > 0.05$, $r = .07$ (see Figure 7.2). However, there was a significant interaction of workload and effort, $F(1.26, 32.99) = 5.04$, $p < 0.01$, $r = .16$, which again highlights the importance of effort under conditions of high workload.

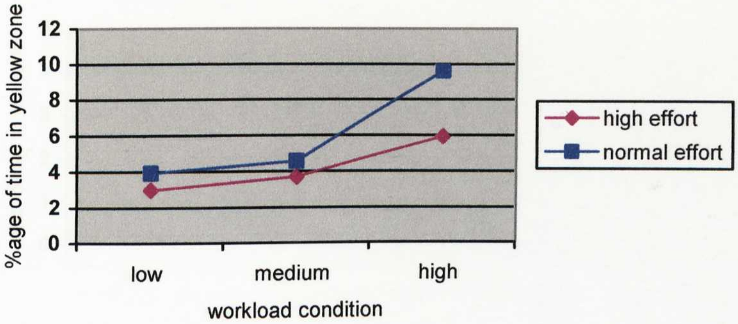


Figure 7.2: Mean percentage of time spent in yellow zone (minor errors)

7.4.1.2 Secondary task performance

As predicted by the Compensatory Control model, primary task performance was largely protected, particularly by operators in the high effort condition. However, following this theoretical model, workload would be predicted to have a greater impact on secondary task performance.

Alarm response time:

Response times, across the entire experimental session, were log transformed and subjected to a 3 x 2 mixed design ANOVA. This surprisingly did not reveal a significant main effect of workload, $F(2, 38) < 1$, $p > 0.05$, $r = .04$, or a significant main effect of effort, $F(1, 19) = 1.41$, $p > 0.05$, $r = .06$, or a significant interaction, $F(2, 38) = 1.80$, $p > 0.05$, $r = .08$. However, with regard to effort, results were in the expected direction, as the high effort group were faster to respond across all three workload conditions (see Figure 7.3). Also, while response times did not change across workload condition for the high effort group, they were marginally slower with increasing workload for the normal effort group. This suggests a trend towards an interaction between workload and effort for this secondary task.

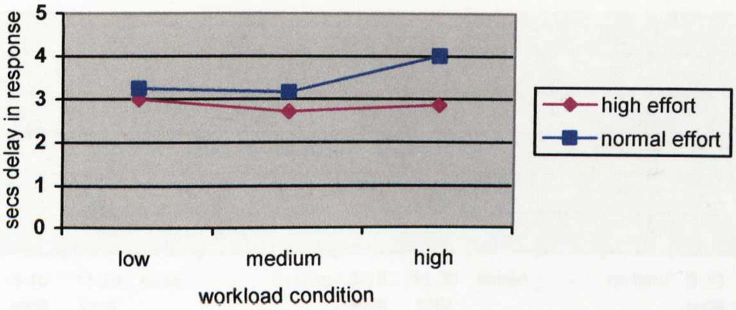


Figure 7.3: Mean alarm response times in seconds

System check performance:

System check performance was calculated in two ways: firstly in terms of the mean time delay to check (number of seconds after designated time); and secondly, in terms of the categorisation of delay as either (1) on-time (0-2 second delay), (2) 'slightly' late (3-10 seconds delay), (3) 'very' late (11-29 seconds delay) and (4) forced system checks (those generated by the system at 30 seconds past the designated time). Figure 7.4 illustrates the patterns of performance for this secondary task for the three workload conditions. Under low and medium workload, the pattern of performance is the same; i.e. participants in both effort conditions most frequently carried out system checks on time, there were fewer 'slightly late' checks, fewer still 'very late' checks and forced system checks were generally uncommon. However, the pattern of performance under high workload was slightly different. In this condition, there were fewer on-time checks, and participants were more frequently 'slightly late' (between 3 and 10 seconds). Furthermore, the mean number of on-time checks was lower for the normal effort group than the high effort group and there were generally a higher number of forced system checks under high workload than low and medium workload.

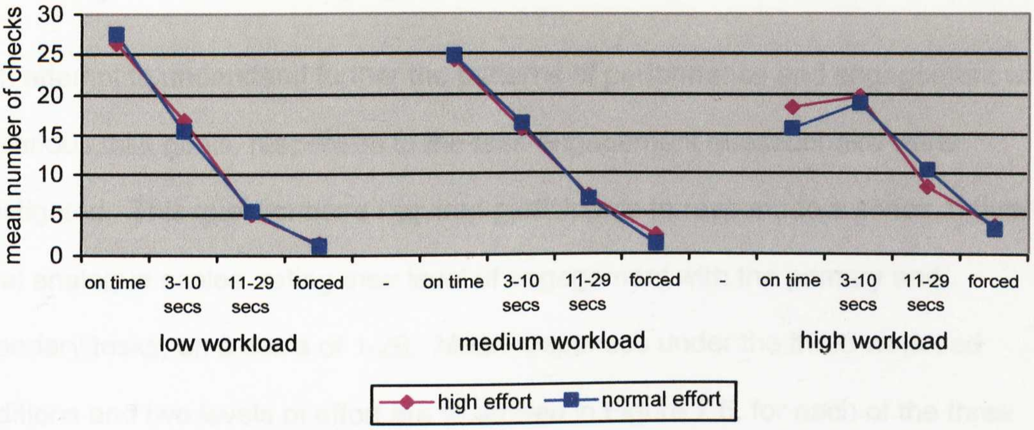


Figure 7.4: Mean number of checks in each category of system check performance

To further investigate this, mean system check delays were log transformed and subjected to a 3 x 2 mixed design ANOVA. Unlike the alarm response secondary task, this analysis revealed a significant main effect of workload, $F(2, 38) = 7.97$, $p < 0.05$, $r = 30$. But, there was no significant main effect of effort, $F(1, 19) <1$, $p > 0.05$, $r = .00$, or significant interaction, $F(2, 38) <1$, $p > 0.05$, $r = .00$. (See Figure 7.5).

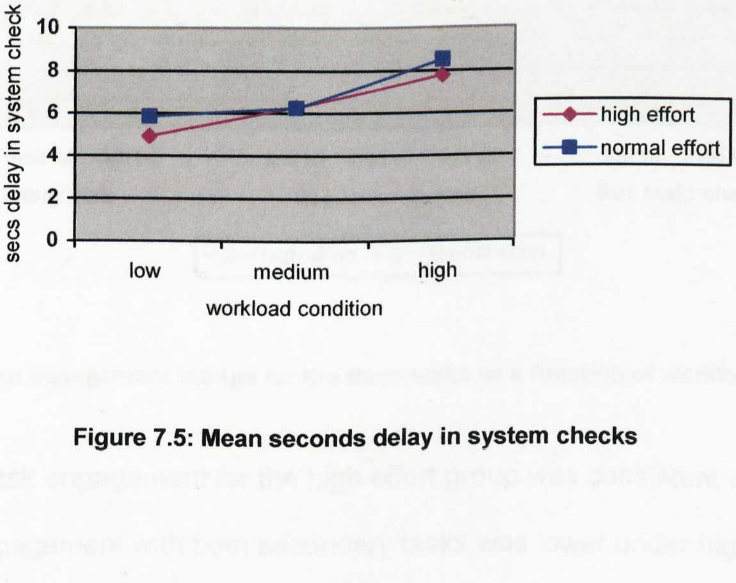


Figure 7.5: Mean seconds delay in system checks

7.4.2 Subjective task engagement

In an attempt to understand further the patterns of performance and engagement with the various task goals, responses to the task-engagement questionnaire were investigated. This questionnaire required participants to respond to a series of three visual analogue scales, rating their level of engagement with the primary and secondary tasks, on a scale of 1-20. Mean responses under the three workload conditions and two levels of effort are illustrated in Figure 7.6, for each of the three tasks.

This figure reveals the different patterns of task engagement for the high and normal effort groups. For the high effort group, engagement with the primary task was consistent across all levels of workload, while for the normal effort group engagement with the primary task increased as a function of workload.

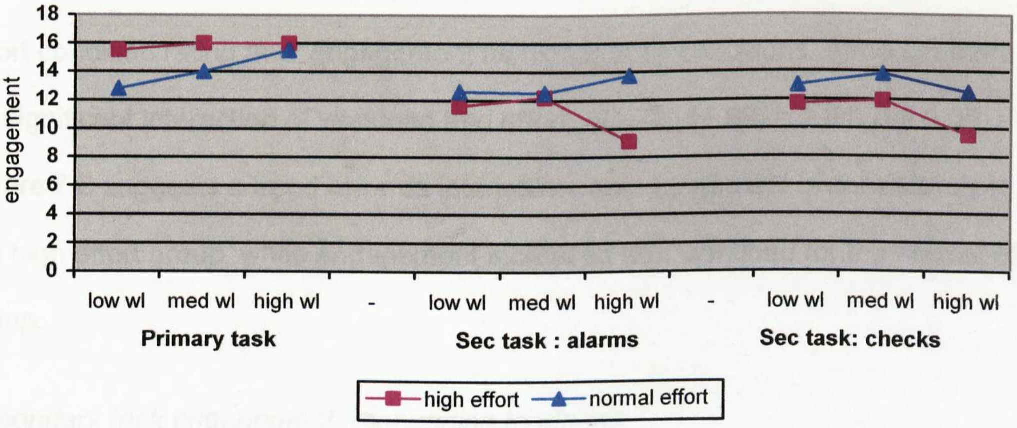


Figure 7.6: Mean engagement ratings for the three tasks as a function of workload and effort

While primary task engagement for the high effort group was consistent across levels of workload, engagement with both secondary tasks was lower under high effort. This suggests a pattern of performance protection for the high effort group.

Ratings of task engagement were less clear for the normal effort group: unlike the high effort group, engagement with the primary task increased with increasing workload; however, ratings of engagement with the secondary tasks showed an increase of engagement under high workload for *alarm response* and decrease of engagement for *system checks*.

To establish the significance of these effects, a series of 3 x 2 mixed design ANOVAs were carried out, one for each of the tasks.

Primary task engagement:

As Mauchly's test for sphericity was found to be significant (X^2 (2) = 7.70, $p < 0.05$), the degrees of freedom were corrected using Huynh-Feldt estimates of sphericity (ε = .86). This analysis revealed a main effect of workload $F(1.73, 41.99) = 3.35$, $p = 0.05$, $r = .11$, with primary task engagement increasing with workload. There was also a significant main effect of effort, $F(1, 26) = 5.37$, $p > 0.05$, $r = .17$, with those in the high effort condition rating their engagement as higher in all conditions. Although there was no significant interaction of workload and effort, $F(1.73, 41.99) = 1.81$, $p > 0.05$, $r = .06$, Figure 7.6 suggests a trend towards interaction, as engagement is consistently high for the high effort group, while engagement increases with workload for the normal effort group.

Secondary task engagement: responding to alarms

The 3 x 2 mixed design ANOVA carried out on engagement ratings for this secondary task did not reveal a main effect of workload $F(2, 52) = 1.13$, $p > 0.05$, $r = .04$, suggesting that increasing workload had no impact on engagement with this secondary task, and there was also no main effect of effort, $F(1, 26) = 2.09$, $p > 0.05$, $r = .07$. However, there was a significant interaction of workload and effort, $F(2, 52) = 6.56$, $p < 0.01$, $r = .20$. As illustrated by Figure 7.6 engagement in this secondary task was consistent for low and medium workload, for both effort groups. However, under high workload, the normal effort group increased their subjective engagement, while the high effort group considerably decreased their engagement.

Secondary task engagement: system checks

Unlike the ANOVA carried out on the first secondary task, the analysis of engagement ratings for system checks did reveal a main effect of workload $F(2, 52) = 3.59$, $p < 0.05$, $r = .12$, suggesting that increasing workload did have an overall impact on engagement with this secondary task (see Figure 7.6). However, there was no main effect of effort, $F(1, 26) = 2.75$, $p > 0.05$, $r = .09$, and no interaction, $F(2, 52) < 1$, $p > 0.05$, $r = .02$.

7.4.3 Subjective workload assessment

To provide a comprehensive picture of all aspects of performance and subjective evaluation, the composite ratings of workload (outlined in section 6.3.5.2) were also subjected to a 3 x 2 mixed design ANOVA. As expected, there was a highly significant main effect of workload $F(2, 52) = 38.27$, $p < 0.001$, $r = .59$, revealing that increasing workload had a major impact on ratings of workload (see Figure 7.7). However, there was no main effect of effort, $F(1, 26) <1$, $p > 0.05$, $r = .01$, and no interaction, $F(2, 52) = 2.20$, $p > 0.05$, $r = .07$.

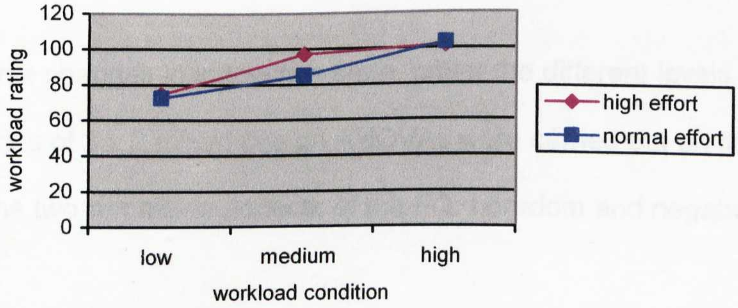


Figure 7.7: Mean workload ratings as a function of workload and effort

7.4.4 Subjective state

As a central component of the fatigue construct is the subjective state of tiredness, it was important to investigate the impact of the experimental manipulations on state fatigue. Having calculated a difference variable for each of the different types of fatigue, changes in mental, physical and sleep-related fatigue were considered. As illustrated in Figure 7.8, the workload manipulation had a differential impact on these three fatigue dimensions. Overall, there were increases in all three types of fatigue under low, medium and high workload. However, while the patterns of change in physical and sleep-related fatigue were similar (with consistent increases throughout the session for all three levels of workload) workload had a different impact on mental fatigue. For this fatigue dimension, the higher the workload, the greater the increase in mental fatigue.

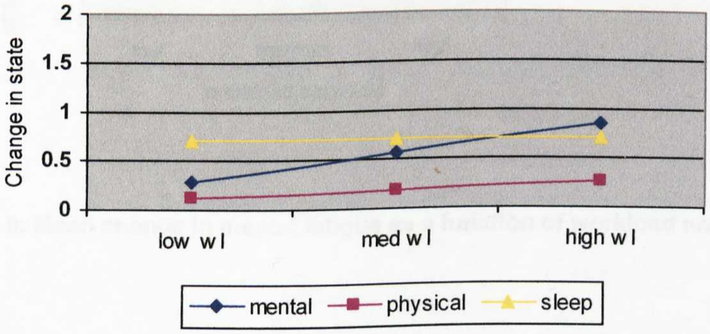


Figure 7.8: Mean change in fatigue dimensions as a function of workload

To further analyse changes in subjective state, under the different levels of workload and effort, a series of 3 x 2 mixed design ANOVAs were carried out on the three types of fatigue and the two remaining aspects of the FQ: boredom and negative affect, *Changes in subjective mental fatigue:*

As expected there was a significant main effect of workload $F(2, 52) = 5.91$, $p < 0.01$, $r = .18$, revealing that increasing workload had a major impact on ratings of mental fatigue. However, while there was no main effect of effort, $F(1, 26) < 1$, $p > 0.05$, $r = .00$, there was a significant interaction, $F(2, 52) = 7.00$, $p < 0.05$, $r = .21$. As illustrated in Figure 7.9, the increases in mental fatigue for the normal effort group were the same under the three levels of workload. However, for the high effort group, there was almost no mean increase in mental fatigue under low workload, and a higher increase in mental fatigue under high workload.

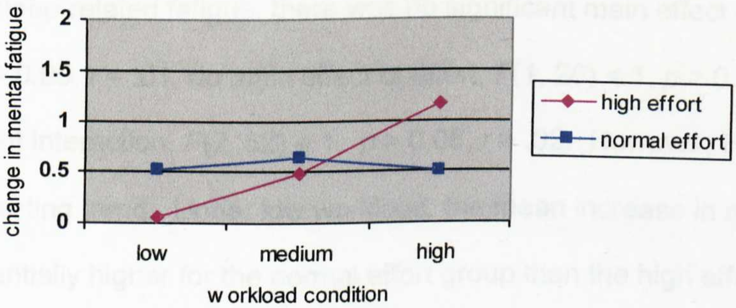


Figure 7.9: Mean change in mental fatigue as a function of workload and effort

Changes in subjective physical fatigue:

Unlike mental fatigue, there was no significant main effect of workload on physical fatigue, $F(2, 52) = 0.44$, $p > 0.05$, $r = .01$, and there was also no main effect of effort, $F(1, 26) <1$, $p > 0.05$, $r = .01$. However, while the interaction between workload and effort failed to reach significance ($F(2, 52) = 2.78$, $p = 0.07$, $r = .09$) p approached alpha and Figure 7.10 illustrates a trend towards an interaction. Although, it important to remember that these changes represent only small mean changes on a 1-5 Likert scale.

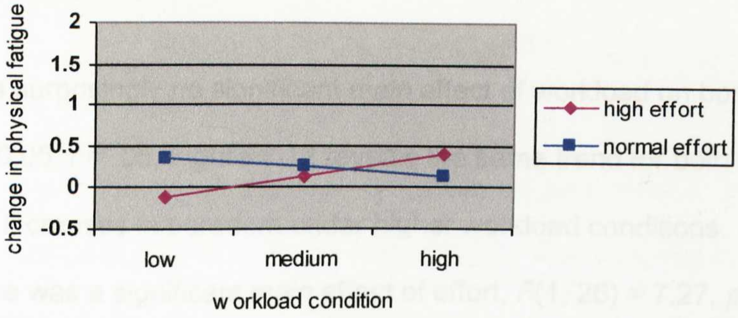


Figure 7.10: Mean change in physical fatigue as a function of workload and effort

Changes in subjective sleep-related fatigue:

With regard to sleep-related fatigue, there was no significant main effect of workload, *F*(2, 52) < 1, $p > 0.05$, $r = .01$, no main effect of effort, *F*(1, 26) < 1, $p > 0.05$, $r = .01$, and no significant interaction, *F*(2, 52) < 1, $p > 0.05$, $r = .02$. However, Figure 7.11 reveals an interesting trend: Under low workload, the mean increase in sleep-related fatigue is substantially higher for the normal effort group than the high effort group.

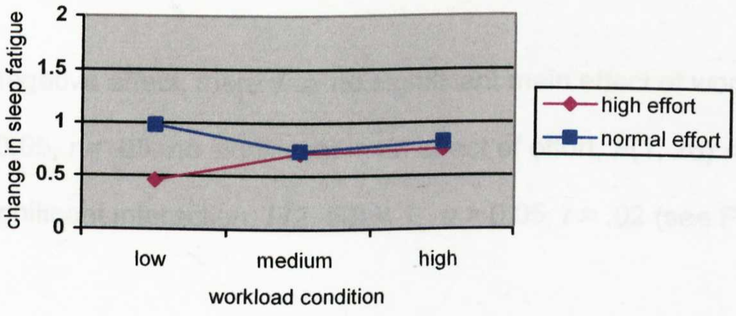


Figure 7.11: Mean change in sleep-related fatigue as a function of workload and effort

Changes in level of boredom:

While there was surprisingly no significant main effect of workload on boredom, *F*(2, 52) =1.59, $p > 0.05$, $r = .05$, Figure 7.12 reveals the same trend for both effort groups: to have smaller increases in boredom under higher workload conditions. As anticipated, there was a significant main effect of effort, *F*(1, 26) = 7.27, $p < 0.01$, $r = .21$. But, there was no significant interaction, *F*(2, 52) <1, $p > 0.05$, $r = .00$.

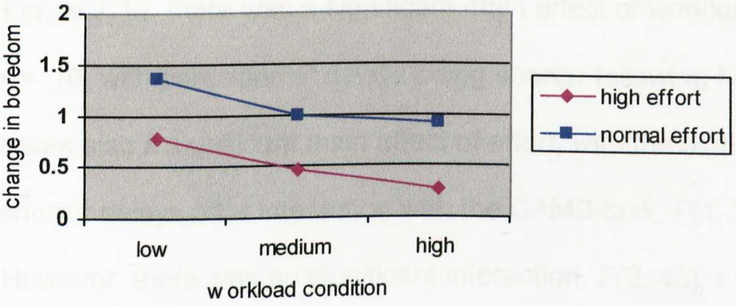


Figure 7.12: Mean change in boredom as a function of workload and effort

Changes in level of negative affect:

With regard to negative affect, there was no significant main effect of workload, *F*(2, 52) =1.48, $p > 0.05$, $r = .05$, no significant main effect of effort, *F*(1, 26) < 1, $p > 0.05$, $r = .01$, and no significant interaction, *F*(2, 52) < 1, $p > 0.05$, $r = .02$ (see Figure 7.13).

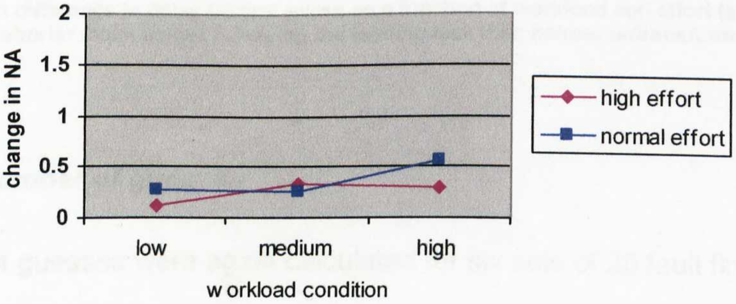


Figure 7.13: Mean change in negative affect as a function of workload and effort

7.4.5 Fatigue after-effects

7.4.5.1 Delay until first guess

Mean delay times until the first guess were calculated for six sets of 25 fault finding trials (pre- and post-CAMS; at three levels of workload). Following log transformation, difference scores were generated by subtracting pre- from post-times. These variables were then subjected to a 3 x 2 mixed ANOVA.

As illustrated in Figure 7.14, there was a significant main effect of workload, $F(2, 46) = 4.61$, $p < 0.01$, $r = .16$, with participants' delays being shorter following higher levels of workload. There was also a significant main effect of effort, with the high effort group tending to yield shorter delays after interaction with the CAMS task, $F(1, 23) = 5.49$, $p < 0.05$, $r = .19$. However, there was no significant interaction, $F(2, 46) < 1$, $p > 0.05$, $r = .03$.

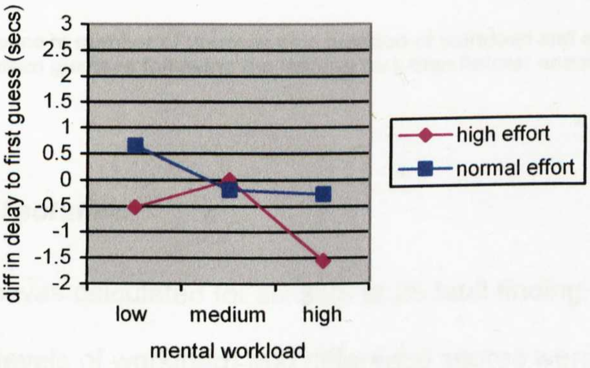


Figure 7.14: Mean difference in delay till first guess as a function of workload and effort (a negative score indicates shorter mean delays following the loading task than before: untransformed data)

7.4.5.2 Mean number of guesses

Mean number of guesses were again calculated for six sets of 25 fault finding trials (pre- and post-CAMS; at three levels of workload) and difference scores were calculated by subtracting log transformed pre- from post-CAMS times. Although

Figure 7.15 reveals a slight trend towards increasing number of guesses as a function of increasing workload, this failed to reach significance in a 3 x 2 mixed ANOVA, *F*(2, 46) = 1.45, $p > 0.05$, $r = .05$. There was, however, a highly significant main effect of effort, with the high effort group tending to make more guesses following each level of workload, while the normal effort group typically made fewer guesses following the CAMS manipulation, *F*(1, 23) = 10.18, $p < 0.01$, $r = .30$. No significant interaction was found, *F*(2, 46) < 1, $p > 0.05$, $r = .00$.

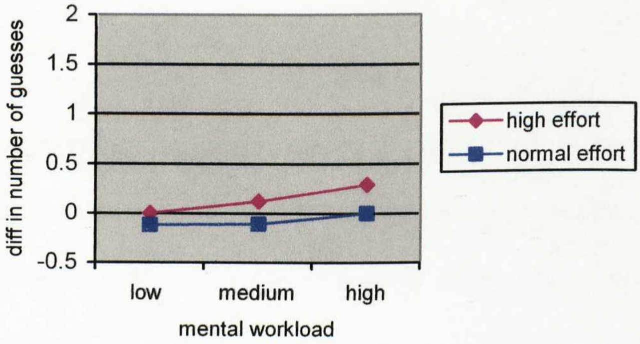


Figure 7.15: Mean difference in number of guesses as a function of workload and effort (a negative score indicates fewer mean guesses following the loading task than before: untransformed data)

7.4.5.3 Mean time till solution

Mean time till solution was calculated for six sets of 25 fault finding trials (pre- and post-CAMS; at three levels of workload) and difference scores were again generated by subtracting log transformed pre- from post-CAMS times. A 3 x 2 mixed ANOVA surprisingly revealed no significant main effect of workload *F*(2, 46) < 1, $p > 0.05$, $r = .00$. However, although the main effect of effort was not found to be significant, there was a trend towards a longer average time to solution for the high effort group, *F*(1, 23) = 3.29, $p = 0.08$, $r = .13$. There was no evidence of an interaction, *F*(2, 46) 1.82, $p > 0.05$, $r = .07$.

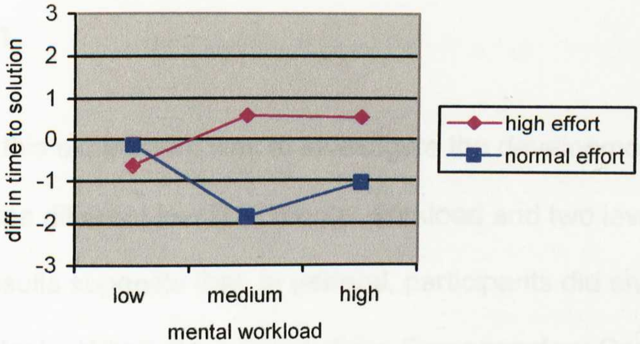


Figure 7.16: Mean difference in time till solution (post minus pre) as a function of workload and effort (a negative score indicates a quicker solution following the loading task than before: untransformed data)

7.5 Discussion

The primary aim of this experiment was to investigate the development and nature of fatigue following three different levels of mental workload and two levels of effort. The overall pattern of results suggests that, in general, participants did engage with the goals of the CAMS task. Within the context of the Compensatory Control model, this was fundamental to the success of the experiment, as one would be unable to attribute any changes in subjective state or performance-related after-effects to fatigue in the absence of effortful task engagement.

Evidence in support of this task engagement was available in terms of CAMS task performance and subjective evaluations. In terms of task performance, there was a generally low level of parameter control failure, suggesting that, even in conditions of high workload, participants were able to protect performance on the primary task. However, of particular interest was the significant interaction of workload and effort on primary task performance, highlighting the importance of high effort for performance protection under conditions of high workload. This was consistent with the subjective ratings of engagement, which were higher for the high effort group with regard to the primary task.

Interestingly, the patterns of performance for the secondary tasks were not quite as expected. It would be predicted by the Compensatory Control model that performance on the primary task would be protected only at a cost to the secondary tasks. However, regarding the secondary task of alarm response, there were no *significant* effects of workload or effort (although there was a trend towards an interaction between workload and effort in the expected direction, for this secondary task). This finding could be interpreted in one of two ways; (1) participants managed to maintain performance across primary and secondary tasks because *even* the high level of workload was not excessive, or (2) participants were sufficiently engaged with this secondary task to also protect performance here. Self report evidence relating to subjective task engagement would suggest the latter was the case, as engagement with both secondary tasks was found to be almost on a par with primary task engagement. Further support for this proposition is provided by the very recent work of Cnossen, Meijman & Rothengatter (2004) who found that performance on secondary tasks while driving was heavily influenced by the nature of the task and its (goal related) proximity to the primary task. In their study, an unrelated working memory task was more affected by higher driving workload than the more closely related secondary task of map reading. This finding offers some explanation for the findings, particularly with regard to the differential sensitivity of the two secondary tasks - it could be argued that the task of alarm acknowledgement is much more closely related to the primary task than the prospective memory task - and only performance on the second secondary task of system checks was affected by workload. Of further interest is the differential performance of the high and normal effort groups on the secondary tasks. The (slightly) superior performance of the high effort group suggests that their protection of primary task performance was not at a cost to secondary tasks. This is consistent with the proposition that this group did in fact recruit additional resources.

The subjective ratings of workload provide further support for the high level of workload associated with managing all system failures. The highly significant difference between ratings suggests a real difference in the subjective experience of the three different levels of workload. Interestingly, there was no significant main effect of effort on subjective ratings of workload, possibly suggesting that ratings of workload were, in this instance, driven more by workload than the level of processing effort actually invested. This is an interesting observation in the light of the comments by Veltman and Gaillard (1996) who highlighted the ambiguity within workload ratings as to whether operators are focusing on whether they *actually* worked hard or thought it would be *necessary* to work hard to perform well.

There are two important points which should be taken from the discussion above: Firstly, that the manipulation of mental workload was successful: As expected from the findings in Experiment One, participants did subjectively experience different levels of workload as intended within the current design. As previously stated, this is extremely important in this study of fatigue to allow the appropriate interpretation of any changes in subjective state or performance after-effects. Secondly, the manipulation of goal commitment via the participant instructions was successful in influencing effort investment. Participants in the high effort group performed better on the primary task under conditions of high workload. This supports the suggestion of Mulder (1986) that effort investment can be directly influenced by several factors including goal commitment.

Regarding changes in subjective states of tiredness, there were increases in mental, physical and sleep-related fatigue under all three levels of workload. Therefore it can be concluded that the operation of CAMS (in the current conditions) generally increased feelings of general tiredness. Of particular note was the different pattern of change for the three FQ dimensions: There were small, but fairly consistent increases in average levels of physical fatigue across low, medium and high workload and, similarly, average increases in sleep-related fatigue were consistent across the three levels of workload (although with regard to sleep-related fatigue the increases were of a higher magnitude). However, average mental fatigue increased with increasing workload.

These different patterns of subjective state change offer some support for a multidimensional state fatigue construct and the distinction between mental, sleep-related and physical state fatigue. (This is consistent with the findings stated in Chapter Four which suggest that individuals are capable of distinguishing, on a subjective level, between different types of fatigue). Most compelling here is the stability of physical and sleep-related fatigue, in comparison to the increased impact of the workload manipulation on mental fatigue. In a broader context, the experience of managing CAMS for 100 minutes is fairly consistent (irrespective of the level of workload). It is possible that taking part in the experimental sessions had a consistent effect on physical and sleep-related fatigue — and that these two distinct states were influenced by the consistent aspects of the task. The only aspect of the task to change was the level of mental workload and, interestingly, the increases in workload were reflected in the increases in subjective mental fatigue. This not only supports the proposition that there is a distinction between the subjective experiences of these types of fatigue, but that they are differentially affected by experiences and conditions. This issue will be discussed in detail in Chapter 10.

The final consideration of the current experiment is the performance after-effects observed on the fault finding task which was carried out prior to, and following, the operation of CAMS. In the context of the theoretical model, the findings, in general, represent a logical progression of the findings discussed above. While the following effects did not all reach significance, a coherent picture was presented suggesting that the delay until the first guess was shorter for the high effort group overall and shorter under conditions of high workload. In the light of the literature described in Chapter 2, this finding is consistent with the premise that fatigued individuals will be more inclined to resort to a low effort strategy (e.g. Holding, 1983). In relation to the Fault Finding task this would result in a quick, less well considered, first guess. While this finding alone could possibly be interpreted that the harder one works, the quicker one is capable of making decisions, performance on the two other Fault Finding task variables are consistent with a low effort strategy under conditions of high workload and high effort. Regarding the average number of guesses until solution, the high effort group tended to make more guesses than the normal effort group, following each level of workload. There was also a small trend towards an increasing number of guesses as a function of increasing workload. (However, in interpreting these findings, it is important to remember that these represent only very small differences in performance). Finally, while there was no overall difference in time until solution for the different levels of workload, the high effort group were typically slightly slower to find a solution following their CAMS operation, whereas the normal effort group were slightly quicker. This again suggests that the high effort group were tending towards quicker, but less accurate, decision-making.

Chapter 8:

Physical Workload and Fatigue

8.1 Summary

The experiment described and discussed in the current chapter aimed to investigate the development of fatigue following different levels of physical workload. While there is a large body of literature relating to the psychological effects of exercise, this work has been extremely varied with regard to its operationalisation of exercise and the psychological variables of interest. A full review of this work is considered to be beyond the scope of this thesis. However, issues relevant to the current programme of research are discussed.

The remainder of the chapter describes and discusses a manipulation of physical workload, with the primary objective of generating a state of 'physical' fatigue. Subjective and physiological indices are utilised to ensure that participants are 'fatigued', as opposed to 'energised'. This is fundamental to the success of this experiment, as the investigation is primarily interested in the impact of physical fatigue on subjective state and performance after-effects.

In summary, both physiological and subjective evidence was found to support the view that participants *were* fatigued following the high level of physical workload. The resulting pattern of subjective state changes was consistent with earlier findings; participants were capable of distinguishing between different types of state fatigue and different conditions give rise to different states. Furthermore, the lack of fatigue-related performance after-effects suggests that after-effects may, to some degree, be specific.

8.2 Introduction

Chapters Six and Seven considered the development of fatigue following the manipulation of different levels of mental workload and effort. Before considering the interaction of multiple sources of fatigue, attention in this chapter is focused on investigating the impact of physical work *alone* on subjective state and fatigue after-effects.

While there is a huge body of literature which has investigated the impact of physical work on psychological variables such as subjective state and cognitive processes, this work has been extremely varied in scope. Although researchers have uniformly attempted to address the seemingly simple issue of whether the effects of exercise are psychologically facilitative or debilitative, the operationalisation of 'exercise' and selection of psychological variables has been so diverse that findings have of course been inconsistent and contradictory. Although a full discussion of this literature is beyond the scope of this thesis, there are a number of issues raised by this work that have some relevance for the central question regarding the nature of fatigue: Essentially, if there are a number of distinct types of fatigue, generated by a range of different conditions, then one could argue that one type of fatigue generator would have no impact on a different type of fatigue. In the context of physical work, excessive levels of exercise should generate a type of fatigue associated with specific subjective feelings and aversion only to further *physical* work. Whereas, if there was only one type of fatigue, generated by different sources, then one might expect physical work to have more generalised after-effects, including some level of aversion to purely cognitive tasks.

Although there is very little research directly considering the impact of physical work on mental fatigue, much of the exercise literature has relevance here, despite addressing a fundamentally different theoretical question.

8.2.1 Exercise and psychological effects

As stated above, exercise and psychological effects have been operationalised in many different ways, leading to contradictory findings. This has limited the development of any coherent theoretical frameworks. However, in an attempt to rectify this, Tomporowski and Ellis (1986) carried out a thorough and systematic review, focusing on the variations and similarities of the various studies and attempting to categorise them on a number of key dimensions: (1) Nature of psychological task; (2) Timing of administration of psychological task; (3) Intensity and duration of physical work; and (4) Prior level of fitness. Within each of these dimensions there are a range of issues relevant to the current study.

8.2.1.1 Nature of the psychological task

It is widely recognised that the nature of the psychological task is an important variable as performance on some tasks appears to be enhanced by exercise, while others appear to be impaired. The types of cognitive task utilised have varied from intellectual tasks, such as complex reasoning (Bills & Stauffacher, 1937; Weingarten & Alexander, 1970) memory-based tasks (Davey, 1973; Sjoberg, 1980; Tomporowski, Ellis & Stevens, 1987) reaction time (Elbel, 1940; Meyers, Zimmerli, Farr & Baschnagel, 1969) and simple mental arithmetic (Gutin & DiGennaro, 1968; Gupta, Sharma & Jaspal, 1974). Therefore, considering the diversity of psychological task selected, it is not surprising that some tasks have been impaired whilst others have shown improvements, or no effect whatsoever. In an attempt to explain these contradictory findings there has been some interesting recent work which emphasises the distinction between stages of information processing, and considers the sensitivity of each stage to the effects of exercise. For example, Fery, Ferry, Vom Hofe, and Rieu (1997) built on the earlier work of Fleury and Bard (1987) and argued that the discrepant results of the vast array of exercise and cognition studies could be explained by the fact that dependent variables such as reaction time are generally utilised as single indices. With regard to fatigue, any sensitivity of the central stage (or decision process) to fatigue could be masked by the combined effects of the preceding (sensory/perceptual) processes and the following (response) processes. Therefore, a dependent variable such as reaction time should be considered as the sum of the effects of physical effort on each stage of cognitive processing. To investigate the selective effects of exercise on the central stage, Fery *et al* used Sternberg's (1969) additive factors method, hypothesising that fatigue could be considered to affect the decision stage if it interacts with the size of the consonant set to be memorised. Interestingly, both aforementioned studies found consistent differential effects of exercise on the separate stages of processing, namely, particular sensitivity of the decision process to fatigue. However, in summarising the limitations of their study, Fery *et al* suggested that future research may benefit from using a concurrent secondary task, such as responding to a tone whilst pedalling, which could have aided the interpretation of the findings because of its sensitivity to any trade-off in attentional resources. This research not only highlights the need to consider the specific stages of information processing, but also illustrates the distinction between the peripheral and central fatigues generated by physical work. The mechanisms of peripheral (muscle) fatigue, widely studied by sports scientists, include specific impairments such as neuromuscular transmission and various metabolic factors that disrupt energy provision and contraction (see Fitts and Metzger, 1983, for a review). As noted by Davis (1995) sports scientists are substantially more interested in peripheral fatigue and typically only attribute a role to central fatigue during exercise when the investigator is unable to find support for a hypothesis at the muscular level, despite the broad recognition among psychologists that the final limit encountered during physical exertion appears to be a psychological rather than a physiological boundary (Holding, 1983). While awareness of the issues associated with peripheral fatigue are clearly of importance in any study manipulating physical work, it is particularly important to note this distinction and the fact that the 'fatigue' referred to in this thesis corresponds to the construct of central fatigue, as categorised by sports scientists, rather than peripheral fatigue.

In terms of the current study, the nature of the after-effects task selected may limit the extent to which it is possible to separate out the effects of physical work on the various stages of information processing. However, awareness of these issues is fundamental in interpreting the findings. Furthermore, following the recommendations of Fery *et al*, it may be important to incorporate a secondary task alongside the physical loading task to aid interpretation of the findings. This would also correspond well with the theoretical framework of the Compensatory Control model - as the physical loading task used here to generate a state of physical fatigue could be considered as a primary task, while a measure of response time to an appropriate stimulus could be considered as a secondary task.

8.2.1.2 Timing of task administration

With regard to timing of administration of the psychological task, existing research has varied as to whether the psychological task is administered during or after the exercise manipulation. Again, considering these findings within the framework of the Compensatory Control model, psychological tasks administered during exercise are less likely to be impaired due to the process of performance protection. Whereas, one might expect performance impairment more likely to manifest after a bout of exercise. However, one would only expect impairment, or fatigue after-effects, when the level of exercise was sufficient to generate a state of fatigue. This would largely be dependent on the following two key dimensions identified by Tomporowski and Ellis: duration and prior level of physical fitness. Without controlling these factors, it is very difficult to interpret the impact of time of test in isolation.

8.2.1.3 Duration and intensity of exercise manipulation

One of the most significant differences identified relates to the *duration* of the exercise. This has varied from single bouts of brief exercise, e.g. Ash (1914); Schwab (1953); Gutin and DiGennaro (1968); Flynn (1972); Gupta, Sharma and Jaspal (1974), to single bouts of more prolonged aerobic exercise, e.g. Bates, Osternig and James (1977); Lichtman and Posner (1983); Gliner, Matsen-Twisdale, Horvath and Maron (1983); McMorris and Graydon (1997), to longer term exposure over periods of weeks or months, e.g. Petruzzello, Landers and Hatfield (1991); Martinsen (1993); Craft and Landers (1998). With regard to longer-term exercise programmes, they have generally been found to have a positive impact on a broad range of "quality of life" factors, such as subjective well-being, energy and depression (see Morgan, 1997, for a review), although their impact on cognitive processes remains debatable (Etnier, Salazar, Landers *et al*, 1997). However, this work is of little relevance here as it essentially focuses on the psychological effects of chronic exercise or physical *fitness*, which is a fundamentally different issue from the current focus of physical fatigue. Much more relevant here are the research findings relating to the acute effects of single bouts of exercise, of varying levels of intensity.

As stated above, one would only expect fatigue after-effects if participants were currently in a state of fatigue. This issue may lie at the heart of the contradictory findings regarding single bouts of exercise as mentally facilitating or debilitating. While many studies have found an enhancement of cognitive performance associated with physical work, there is no shortage of studies finding either no effect or substantial impairments. Furthermore, a number of studies have found an inverted U relationship between exercise and cognition: Davey (1973), the most frequently quoted of these, found subjects' performance on a short-term memory detection task improved after two minutes of stationary cycling, but became impaired after 10 minutes. Similarly, Gupta, Sharma and Jaspal (1974) also found an initial facilitation of simple arithmetic following two and five minutes of step-ups, but noted an impairment of the task after ten and fifteen minutes.

These findings are consistent with the inverted U hypothesis first proposed by Yerkes and Dodson (1908) to explain the relationship between arousal and psycho-motor task performance. Such theories suggest that as physical arousal increases, performance will improve up to a point and then deteriorate with further increases in arousal. While theories of arousal have generally been criticised for their limited explanatory value, a notable theory is that of Easterbrook (1959) whose Cue Utilisation theory explains the effect of physical arousal on performance in terms of narrowing of attention. This theory, consistent with the inverted U hypothesis, proposes that variations in physical arousal will produce a change in attentional processes. Specifically, increases in arousal results in a narrowing of attention to those task components which are important for performance. Attention to peripheral information, which could distract attention, is reduced. Then, if the level of physical arousal increases, the continued narrowing of attention may inhibit the selection of task-relevant cues. This theory was particularly influential and there is currently a great deal of evidence to support the proposition that attentional capacity varies during the performance of cognitive tasks and that potential capacity is affected by physical arousal levels (Kahneman, 1973). Such theories clearly offer an explanation for the facilitative effects of single bouts of exercise on cognitive tasks, and it is not difficult to accept that physical activity, sufficient to increase physiological activation, could enhance performance up to a point. However, as the intensity and duration of the physical work increases, there will be a point at which performance is expected to deteriorate. To what extent this deterioration can be attributed to cognitive processes such as attentional narrowing, or the (essentially motivational) state of fatigue, is both extremely interesting and highly debatable. The possibility that exercise does not alter cognitive functioning, but that it is motivational variables which affect task performance is a hypothesis put forward by Tomporowski and Ellis (1986). This issue will be returned to in Chapter 10.

With regard to the current study, the generation of a state of fatigue through physical work is the primary aim, rather than an investigation of the psychological or cognitive effects of exercise per se. Therefore, it is vital that the manipulation of physical work is sufficiently intense to ensure participants are genuinely fatigued, rather than being 'activated' by the exercise. This will clearly interact with the final dimension: level of participant fitness.

8.2.1.4 Prior level of participant fitness

The final dimension of interest is that of physical fitness. While a large proportion of research programmes carry out no preliminary index of fitness (Davey, 1973; Gupta, Sharma & Jaspal, 1974) others have assessed maximal strength (Stauffacher, 1937; Andreassi, 1965) or aerobic capacity (Sjoberg, 1980; Weingarten & Alexander, 1970; Tomporowski *et al*, 1985). The impact of fitness as a factor is not unequivocal, however in the cases of Sjoberg (1980) and Weingarten and Alexander (1970) there were impairments of performance only in the case of the low fitness participants. While it could be argued that this occurred because only the low fitness participants were truly fatigued, this is not possible to determine in retrospect.

8.2.2 Experiment Three rationale

As discussed above, there is a large literature relating to exercise and mental performance. However, the emphasis of the current study is on investigating the dynamic process of fatigue following different levels of physical work. Physical work will be manipulated via different intensities of stationary cycling; all participants in all conditions being required to maintain a constant speed, but working at different resistance levels.

However, to investigate the effects of controlling the required pace, an additional variable of pace control will be manipulated, with participants being required to control the pace either very closely, or fairly flexibly.

Therefore, the current experiment aims to investigate the development of fatigue following different levels of physical workload, under different levels of pace control.

8.3 Method

8.3.1 Design

A mixed design was used with two independent variables: (1) Physical workload and (2) Pace-control. Physical workload was varied at three levels; low, medium and high, determined by the level of resistance programmed into a stationary cycle. This IV was manipulated within subjects, resulting in three experimental sessions. Pace-control was manipulated between subjects, on two levels; high and low control.

Each experimental session lasted approximately one hour: Participants were required to complete 25 Fault Finding networks before and after 30 minutes of stationary cycling.

To minimise any carryover effects, a minimum delay of three days between experimental sessions was enforced. The maximum delay was seven days.

8.3.2 Participants

For this study, participants were recruited through advertisements placed in the Sports Centre at the University of Hull. Twenty-six students initially volunteered to participate in the experiment and were invited to take part in 'an experiment investigating the effects of physical exercise on performance'. A prerequisite for participation was a good level of general fitness and regular exercise. All volunteers exercised, on average, between 1 hour and 5 hours per week (M = 2.39 hours, SD = 1.01 hours). However, before volunteers were accepted as participants, they were required to undergo an assessment of their maximum oxygen uptake (VO_2 max) as an indicator of cardio-respiratory fitness (Astrand & Rodahl, 1986: See section 8.3.3.1). This process, carried out by a physiologist, performed two functions: Firstly, the level of VO_2 max was used to ascertain suitability for the experiment; secondly, this measure of fitness was used to determine individual levels of workload, relative to the fitness of the individual, rather than set at an absolute level of demand (discussed in more detail below). Two volunteers were considered to be not sufficiently fit to take part in the experiment: One suffered knee pains as a result of the procedure and one failed to meet the minimum VO_2 max criteria set by the physiologist.

The final experimental sample of 24 consisted of 15 men and 9 women, aged 18-36 (mean = 24.35, SD = 3.32).

8.3.3 Experimental Procedure

8.3.3.1 Pre-experimental fitness testing

VO_2 max testing was used to assess the participants' general physical fitness. This recognised procedure assesses the maximum oxygen uptake, breathing at sea level (Astrand & Rodahl, 1986). Maximum oxygen consumption was estimated from performance on standardised protocols of incremental exercise to volitional exhaustion. Exercise is most frequently manipulated either by running or cycling. For the fitness testing and the experimental sessions, cycling was adopted as the exercise task, as this provided options for carrying out computer based tasks concurrently. A static bicycle ergometer (Monarch 824E) was used, which allows the resistance of the pedals to be manipulated via weights added to a weight basket.

The protocol requires the participant to begin pedalling at a constant speed of 60 RPM. During the first three minutes, they are required to pedal at a low resistance as a warm-up. Following this, the load is increased by 0.5 kg every three minutes, until the participants can no longer pedal.

8.3.3.2 Pre-experimental training

Twenty-four of the 26 participants who took part in the fitness testing progressed to the next stage. The training principles for the Fault Finding task adopted for Experiment 2 were maintained, i.e. participants were trained in small groups, with sole access to a PC. However, as the Fault Finding task was the only computer program which required training, subjects were provided with one training session, lasting one hour.

Participants were provided with information about the aims and procedures of the current experiment and were then introduced to the Fault Finding task. This included a description of the basic features and operating principles, which was followed by an explanation of the possible strategies available to the participants (see section 5.6.3). Finally, participants were asked to work through a series of 75 networks, employing all potential strategies to meet the dual goals of speed and parsimony, which were not prioritised. Performance on this task was again monitored to ensure full understanding of the task and to minimise the possibility of learning within the experimental sessions.

8.3.3.3 Experimental Procedure

The experimental procedure of previous experiments was straightforward, involving only paper and pencil questionnaires and interaction with the two computer programs. However, the procedure for this experiment was more complex because of the involvement of physical workload and physiological measures, alongside the computer-based tasks. Therefore, the experimental procedure will be described below.

For both practical and theoretical reasons, it was considered important that participants would carry out the computer based tasks while in the same physical 'situation' as the physical loading task: In practical terms, Experiment 4 would manipulate concurrent physical and mental workloads, requiring participants to cycle whilst operating the computer simulator. In theoretical terms, fatigue after-effects are arguably more likely to manifest if participants are not allowed to remove themselves from their 'working' position before completing the after-effects task. Therefore, a computer stand was designed which was fitted around the front of the stationary cycle, positioning the computer monitor on a platform set approximately at eye level. This stand also incorporated a moveable platform which could be lowered when using the keyboard/mouse or when completing paper and pencil questionnaires, and raised during the 30 minutes of physical exercise, when the computer was not in use.

Before starting the session, the participants were fitted with a heart rate monitor, including a chest strap and 'watch'. Participants were then invited to adjust the seating height and handlebars of the cycle to a comfortable position. (This position was recorded at the first session and maintained for the following two sessions to remove any potential impact of different seating positions.) Once participants were seated on the cycle, the experimenter explained the broad aims of the experiment and outlined the procedure. Participants were then asked to relax for a period of two minutes, in order to provide a baseline measure of resting heart rate.

The participants were then asked to complete 25 Fault Finding networks, followed by the state fatigue and state anxiety questionnaires. At this point, participants were fitted with a reaction time device (described in section 8.3.5.1 below) and given 5 practice trials of this task, to familiarise themselves with it.

Participants were then required to cycle for 30 minutes under one of three conditions of physical workload, which was followed by the completion of the state questionnaires and, finally, 25 additional Fault Finding networks.

8.3.4 Independent Variables

This study involved the manipulation of two IVs; physical workload and effort.

Physical workload was varied at three levels; low, medium and high, manipulated by varying the intensity of the physical demands. This was achieved by maintaining a speed of 60 RPM at all workload levels, while adjusting the resistance to pedal rotation by adding the appropriate weight to the cycle. The high workload condition required participants to work for 30 minutes at 70% of their maximum capacity, medium workload was 50% and low workload required the participants to maintain a speed of 60 RPM with minimal weight resistance (only the weight of the basket). Participants were randomly allocated to one of six possible experimental schedules, which determined the order in which they experienced the three workload conditions.

Once assigned to a workload schedule, each participant was randomly allocated to one of two pace-control conditions.

Pace-control was varied at two levels; controlled pacing and flexible pacing. (1) In the controlled pacing condition participants were required to strictly monitor the RPM and adjust the rate of cycling as soon as it deviated from the recommended 60 RPM. (2) The flexible pacing condition required participants to keep the rate of cycling at approximately 60 RPM with occasional monitoring 'but not to be too concerned if this deviates about 5 RPM above or below'.

8.3.5 Dependent Variables

8.3.5.1 Performance-related dependent variables

Secondary task: Simple reaction time

As previously mentioned, this experiment incorporated a simple reaction time task. In accordance with the Compensatory Control model, the 30 minutes of cycling (either *at* or *around* 60 RPM) can be considered to be a primary task. However, it was argued to be worthwhile to introduce a secondary task to the current experiment, as a measure of latent degradation. As previously discussed, reaction time has frequently and successfully been utilised as a secondary task, and hence was adopted here (although the questionable ecological validity is recognised). To avoid interference with the task of pace-control, it was decided that an auditory reaction time task would be more appropriate than a visual based task. Therefore, the secondary task involved an auditory alarm which sounded at random intervals, with one alarm in each 1.5 minute period. This resulted in approximately 20 alarms during the cycling phase of the experimental session. Participants were required to respond, as quickly as possible, by pinching forefinger and thumb together. The reaction time was measured by a recording device which initiated a timer as the alarm sounded. This was stopped with the connection of two electrodes attached to the forefinger and thumb of the dominant hand.

Fault Finding task performance

As with Experiment 2, participants were required to complete 25 Fault Finding networks before and after the cycling. This provided the same series of dependent variables, namely mean and median measures of (1) time until first guess (2) time until solution and (3) number of guesses.

8.3.5.2 Subjective measures

As with the previous experiment, state fatigue was measured with the 15 item Feelings Questionnaire, state anxiety was measured using the STAI Form Y-2 and the Workload Questionnaire was used to evaluate the nature of the demands associated with the task.

An additional measure for the current experiment was a further workload questionnaire, based on the Borg scale (1978). This required participants to rate *physical demands, mental demands* and *effort* on a 20 point scale, with verbal anchors of 'very low' (1) and 'very high' (20). This scale was presented at 3 intervals during the cycling: after 10 minutes, 20 minutes and 29 minutes. Participants were given the option of responding verbally, by saying a number between 1 and 20, or by pointing at the number which best reflected their rating.

8.3.5.3 Physiological measures

Heart rate was measured throughout the entire experimental session via a radiotelemetry pulse monitor (Polar Electro OY). Participants were required to wear a chest strap, which recorded cardiac activity and transmitted the data to a pulse-monitor worn on the wrist. Following each experimental session, heart rate data was downloaded onto a laptop computer.

8.3.6 Data treatment

As with the previous two experiments, Analysis of Variance was again the main method of data analysis. Continuing with the principles of data treatment where sphericity has been violated, degrees of freedom were corrected using either Greenhouse-Geisser (< 0.75) or Huynh-Feldt (> 0.75) estimates of sphericity, in accordance with Field (2000).

Heart rate data was essentially utilised in the current experiment to confirm that participants in the low, medium and high physical workload conditions were working at significantly different physiological intensities. Therefore, analysis of the physiological data was limited to simple measures of heart rate (average number of beats per minute). This data was broken down into three time periods of (1) heart rate during the Fault Finding task – prior to cycling; (2) heart rate during cycling; (3) heart rate during the Fault Finding task – following the cycling

The auditory response time data was found to be positively skewed. Therefore, as with previous performance measures (in Experiments One and Two) they were log transformed prior to any analysis. However, data presented in any Tables or Figures reflect untransformed data. To analyse this dependent variable further, the data was split into two time periods for each cycling session, to investigate any effects of time on task. This provided a second repeated measures variable for the mixed design ANOVA.

The data from the Fault Finding Task was extracted using the Data Extraction programme and entered into SPSS. This data was again log transformed to reduce skew, before the calculation of difference variables.

The questionnaire data was again reduced into their relevant factors: The workload questionnaire was reduced into a single index of overall workload; state fatigue was reduced into five factors of mental, physical and sleep-related fatigue, boredom and negative affect, and difference scores were then obtained by subtracting pre- from post-factor scores. Data from the state anxiety questionnaire was reversed where necessary and reduced into a single index score.

8.4 Results

8.4.1 Physiological data: Heart rate

To investigate the impact of the three levels of physical workload on heart rate, this data was submitted to a 3 (workload) x 3 (time period) x 2 (pace control) mixed design ANOVA.

As anticipated, this analysis revealed a highly significant main effect of workload, *F*(2, 38) = 150.03, $p < 0.001$, $r = .88$, with heart rate increasing with increasing physical load (see Figure 8.1). To confirm that each level of workload was significantly different from the adjacent level, the data was then submitted to a Bonferroni post-hoc test. This revealed a highly significant difference between all levels of physical workload (low - medium workload, mean difference = -24.71, $p < 0.001$; medium - high workload, mean difference = -12.32, $p < 0.001$).

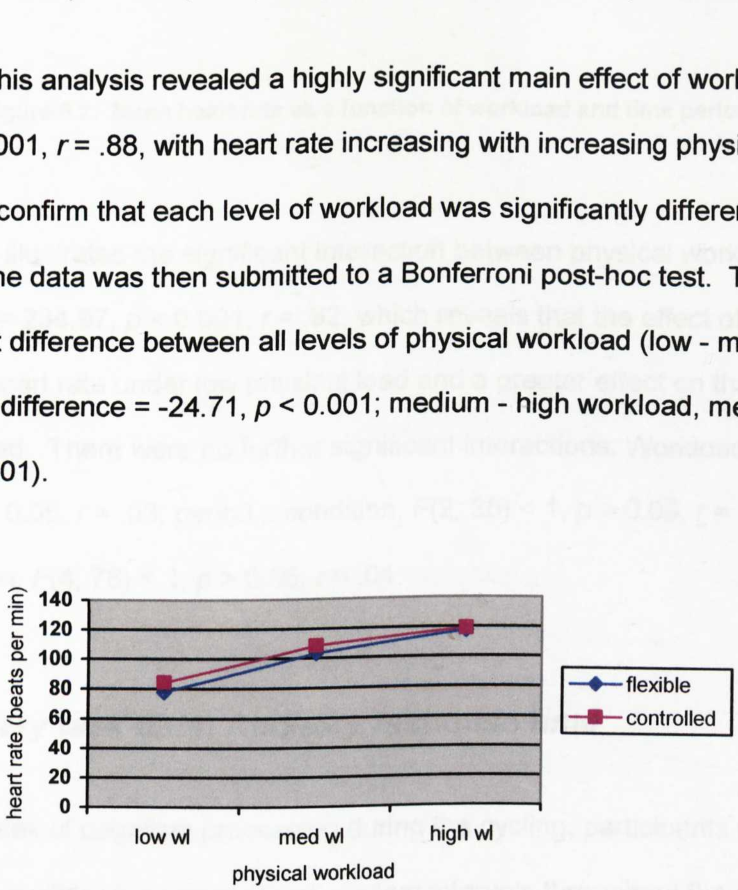


Figure 8.1: Mean heart rate as a function of workload and pace control

As also suggested by Figure 8.1, there was no significant main effect of pace control, *F*(1, 19) < 1, $p > 0.05$, $r = .03$. However, as anticipated, there was a highly significant main effect of time period, as participants' heart rates increased in response to the cycling in period two, *F*(2, 38) = 555.00, $p < 0.001$, $r = .96$. (See Figure 8.2.)

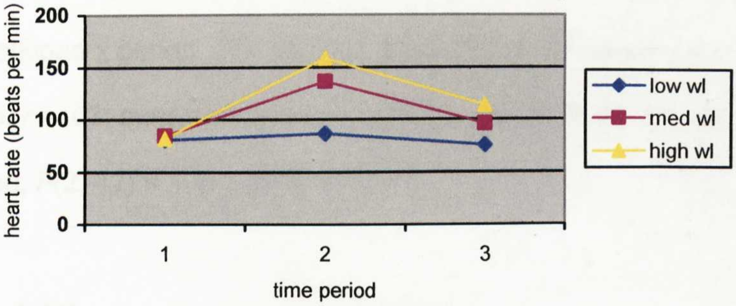


Figure 8.2: Mean heart rate as a function of workload and time period

This Figure also illustrates the significant interaction between physical workload and time period, *F*(4, 46) = 234.97, $p < 0.001$, *r* = .92, which reveals that the effect of period had little impact on heart rate under low physical load and a greater effect on those in medium and high workload. There were no further significant interactions: Workload x condition, *F*(2, 38) < 1, $p > 0.05$, *r* = .03; period x condition, *F*(2, 38) < 1, $p > 0.05$, *r* = .00; workload x period x condition, *F*(4, 76) < 1, $p > 0.05$, *r* = .01.

8.4.2 Secondary task data: Auditory response time

To provide an index of cognitive processing during the cycling, participants were required to respond to an auditory tone sounded at random intervals throughout the cycling session. Response times were again log transformed and subjected to a 3 x 2 x 2 mixed design ANOVA.

As anticipated, this analysis did reveal a significant main effect of workload, *F*(2, 42) = 6.77, $p < 0.01$, *r* = .24, with response time increasing with increasing physical load (see Figure 8.3). Figure 8.3 also suggests a trend towards an effect of time period, as participants were typically faster to respond in the first fifteen minute period. However, this effect failed to reach significance, *F*(1, 21) = 1.34, $p > 0.05$, *r* = .06. There was also no significant main effect of pace control, $F(1, 21) < 1$, $p > 0.05$, $r = .00$, and no significant interactions: Workload x period, $F(2, 42) < 1$, $p > 0.05$, $r = .01$; workload x condition, $F(2, 42) < 1$, $p > 0.05$, $r = .03$; period x condition, $F(2, 21) < 1$, $p > 0.05$, $r = .00$; workload x period x condition, $F(2, 42) < 1$, $p > 0.05$, $r = .03$.

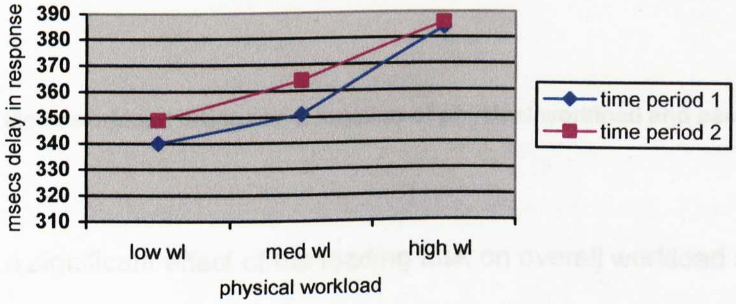


Figure 8.3: Mean alarm response times in milliseconds for the two time periods as a function of workload and pace control

8.4.3 Subjective post-task workload assessment

8.4.3.1 Overall post-task workload assessment

To investigate subjective evaluations of the loading task (cycling at a specified rate and responding to auditory tones), the composite ratings of workload (outlined in section 6.3.5.3) were subjected to a 3 x 2 mixed design ANOVA. As expected there was a highly significant main effect of workload $F(2, 42) = 64.23$, $p < 0.001$, $r = .75$, revealing that increasing physical load had a major impact on overall ratings of workload (see Figure 8.4). However, there was no main effect of pace control, $F(1, 21) < 1$, $p > 0.05$, $r = .00$, and no interaction, $F(2, 42) = < 1$, $p > 0.05$, $r = .04$.

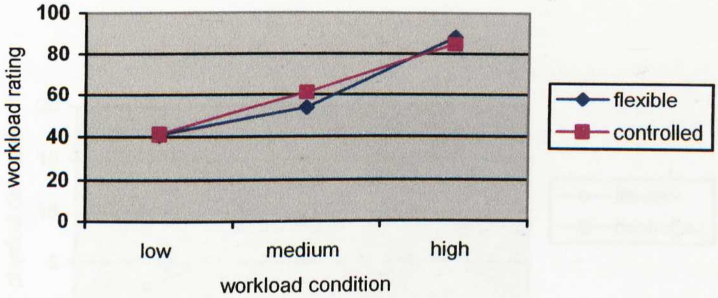


Figure 8.4: Mean workload ratings as a function of physical workload and pace control

Having identified a significant effect of the loading task on overall workload assessment, it is relevant to consider the contribution of separate items to this effect. Of particular interest are the physical and attentional demand items, and the effect of workload and pace control on these individual items.

8.4.3.2 Workload item assessment: Physical demands

Ratings of physical demands were subjected to a 3 x 2 mixed design ANOVA. As expected there was a highly significant main effect of workload $F(2, 42) = 155.06$, $p < 0.001$, $r = .88$, revealing that increasing physical load had a major impact on ratings of physical demand (see Figure 8.5). However, there was no main effect of pace control, $F(1, 21) < 1$, $p > 0.05$, $r = .00$, and no interaction, $F(2, 42) < 1$, $p > 0.05$, $r = .04$.

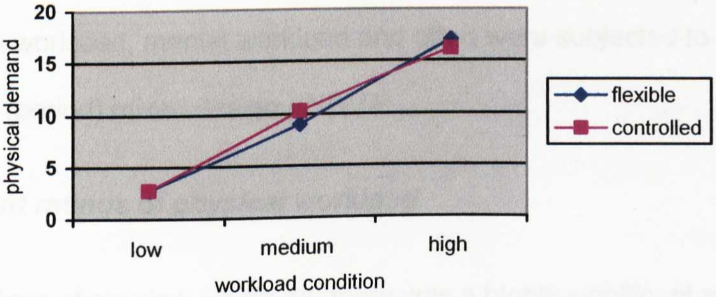


Figure 8.5: Mean physical demand ratings as a function of physical workload and pace control

8.4.3.3 Workload item assessment: Attentional demands

Ratings of attentional demands were also subjected to a 3 x 2 mixed design ANOVA and there was again a highly significant main effect of workload $F(2, 42) = 8.40$, $p < 0.001$, $r = .28$, which revealed that increased physical load had an impact on ratings of attentional demand (see Figure 8.6). However, while the attentional demands were rated higher in the controlled condition (particularly under low workload) there was surprisingly no significant main effect of pace control, $F(1, 21) < 1$, $p > 0.05$, $r = .03$, and there was also no significant interaction, $F(2, 42) < 1$, $p > 0.05$, $r = .02$.

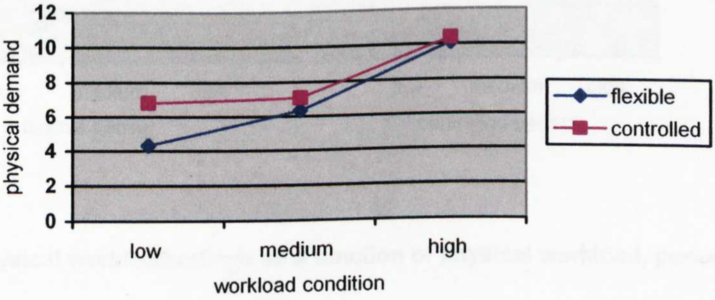


Figure 8.6: Mean attentional demand ratings as a function of physical workload and pace control

8.4.4 Subjective mid-task workload assessment

To investigate further subjective evaluations of the physical loading task, the concurrent ratings of physical workload, mental workload and effort were subjected to a 3 (wl) x 2 (pace control) x 3 (period) mixed design ANOVA.

8.4.4.1 Concurrent ratings of physical workload

With regard to ratings of physical workload, there was a highly significant main effect of physical workload, $F(2, 44) = 127.21$, $p < 0.001$, $r = .85$, revealing that increasing physical load had a major impact on ratings of physical workload. There was a main effect of period $F(2, 44) = 59.57$, $p < 0.001$, $r = .73$, revealing that ratings of physical workload increased throughout the session, but there was no main effect of pace control, $F(1,21) < 1$, $p > 0.05$, $r = .01$. With regard to interactions, there was a significant workload x period interaction $F(4, 88) = 23.93$, $p < 0.001$, $r = .52$, revealing that ratings increased more over the session, in the higher workload conditions (see Figure 8.7). However, there were no further interactions: workload x condition, $F(2, 44) < 1$, $p > 0.05$, $r = .02$; period x condition, $F(2,44) < 1$, $p > 0.05$, $r = .01$; workload x period x condition, $F(4, 88) < 1$, $p > 0.05$, $r = .03$.

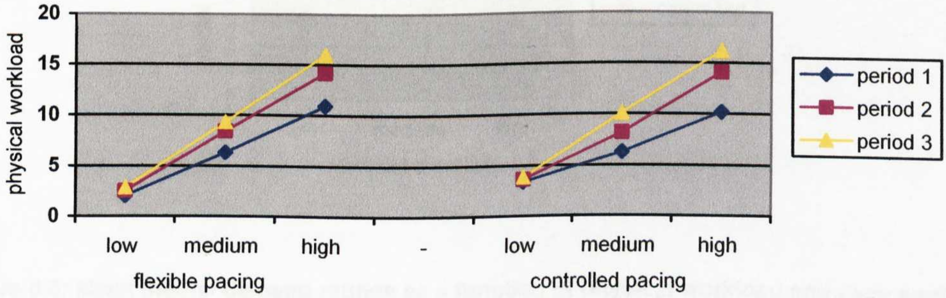


Figure 8.7: Mean physical workload ratings as a function of physical workload, period and pace control

8.4.4.2 Concurrent ratings of mental demand

The ratings of mental demand throughout the cycling session were also subjected to a 3 x 2 x 3 mixed design ANOVA. Mauchly's test of sphericity was found to be significant in all three effects: workload (X^2 (2) = 15.90, p < 0.001); period (X^2 (2) = 20.92, p < 0.001); workload x period (X^2 (9) = 35.83, p < 0.001) . Therefore, in each case, the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε = .64, .61, .58 respectively). There was again a highly significant main effect of workload $F(1.30, 28.73) = 20.09$, $p < 0.001$, $r = .47$, revealing that increasing physical load had a major impact on ratings of mental demand. There was also a main effect of period $F(1.22, 28.73) = 14.51$, $p < 0.001$, $r = .39$, revealing that ratings of mental workload also increased throughout the session. But there was no main effect of pace control, $F(1, 22) < 1$, $p > 0.05$, $r = .01$. As illustrated in Figure 8.8, mental workload was rated higher for the controlled pace condition under each level of workload, although these differences were smaller than anticipated.

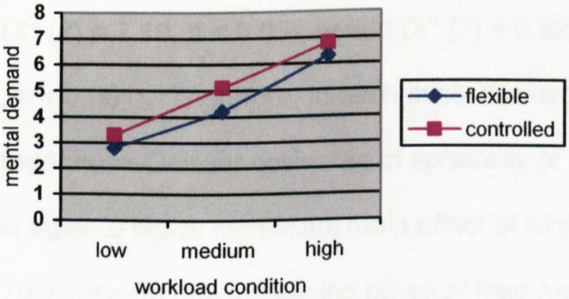


Figure 8.8: Mean mental demand ratings as a function of physical workload and pace control

With regard to interactions, there was a significant workload x period interaction $F(2.35, 51.81) = 8.76$, $p < 0.001$, $r = .28$, revealing that mental workload ratings increased more over the session, in the higher workload conditions (see Figure 8.9). However, there were no further interactions: workload x condition, *F*(2, 44) < 1, $p > 0.05$, $r = .00$; period x condition, *F*(2,44) < 1, $p > 0.05$, $r = .01$; workload x period x condition, *F*(4, 88) < 1, $p > 0.05$, $r = .03$. (See Figure 8.9.)

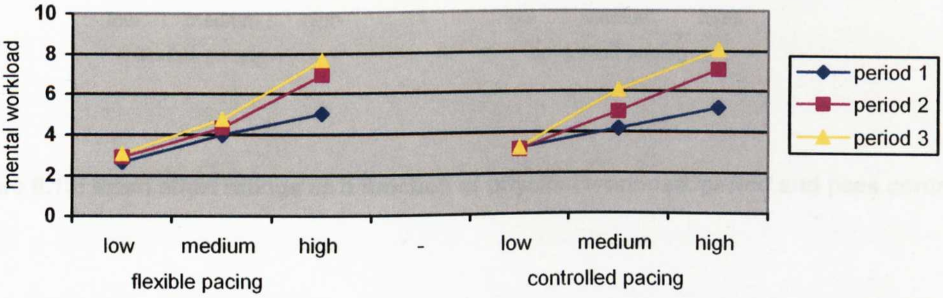


Figure 8.9: Mean mental workload ratings as a function of physical workload, period and pace control

8.4.4.3 Concurrent ratings of effort

The ratings of effort throughout the cycling session were also subjected to a 3 x 2 x 3 mixed design ANOVA. Mauchly's test of sphericity was again found to be significant in all three effects: workload (X^2 (2) = 7.19, $p < 0.05$); period (X^2 (2) = 9.62, $p < 0.01$); workload x period (X^2 (9) = 28.79, $p < 0.001$). Therefore, in each case, the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .75, .73, .58$ respectively). There was again a highly significant main effect of workload *F*(1.55, 34.10) = 72.88, $p < 0.001$, $r = .76$, revealing that increasing physical load had a major impact on ratings of effort. There was also a main effect of period *F*(1.46, 3.23) = 33.90, $p < 0.001$, $r = .60$, revealing that ratings of effort also increased throughout the session, but there was again no main effect of pace control, *F*(1, 22) < 1, $p > 0.05$, $r = .00$. With regard to interactions, there was a significant workload x period interaction *F*(2.32, 51.11) = 19.61, $p < 0.001$, $r = .47$, revealing that ratings increased more over the session, in the higher workload conditions (see Figure 8.10).

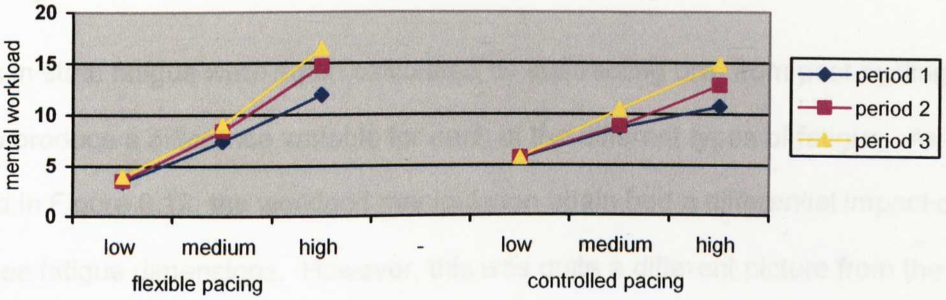


Figure 8.10: Mean effort ratings as a function of physical workload, period and pace control

However, unlike ratings of mental and physical workload, there was also a significant interaction between workload and pace control, *F*(2, 44) = 3.24, $p < 0.05$, *r* = .12, revealing that under low and medium load those in the controlled pace condition reported higher levels of effort. Whereas, in the high workload condition, those in the flexible group reported higher ratings of effort (see Figure 8.11). However, this was only a small effect. There were no further interactions: period x pace control, *F*(2, 44) < 1, $p > 0.05$, *r* = .02; workload x period x pace control, *F*(4, 88) = < 1, $p > 0.05$, *r* = .03.

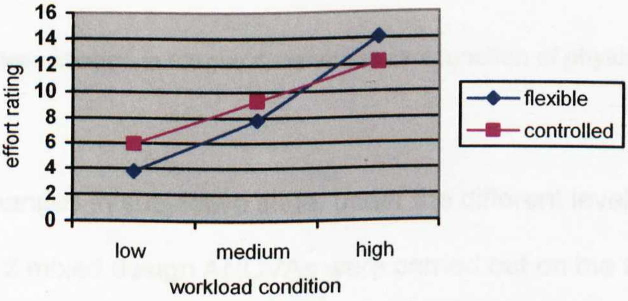


Figure 8.11: Mean effort ratings as a function of physical workload and pace control

8.4.5 Subjective state

Changes in state fatigue were again calculated by subtracting pre- from post-cycling scores to produce a difference variable for each of the different types of fatigue. As illustrated in Figure 8.12, the workload manipulation again had a differential impact on these three fatigue dimensions. However, this was quite a different picture from the changes in fatigue following mental work. Overall, there were very small average decreases in mental and sleep-related fatigue, and only under high workload was there a very slight increase in average mental fatigue. As one would anticipate, the pattern for physical fatigue was different: There was a decrease in physical fatigue under low workload and linear increases under medium and high workload.

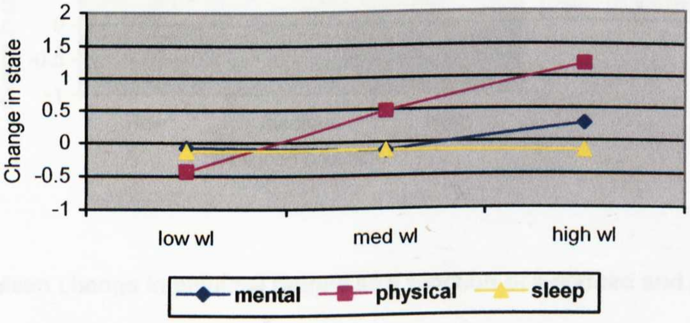


Figure 8.12: Mean change in fatigue dimensions as a function of physical workload

To further analyse changes in subjective state, under the different levels of workload and effort, a series of 3 x 2 mixed design ANOVAs were carried out on the three types of fatigue and the two remaining aspects of the subjective state measure: boredom and negative affect.

Changes in subjective physical fatigue:

As Mauchly's test of sphericity was found to be significant (X^2 (2) = 6.90, $p < 0.05$), the degrees of freedom were corrected using Huynh-Feldt estimates of sphericity (ε = .86). As anticipated, this analysis revealed a highly significant main effect of physical workload on physical fatigue, *F*(1.73, 38.20) = 16.10, $p < 0.001$, *r* = .42, but despite those in the flexible condition reporting higher increases in physical fatigue, this difference failed to reach significance, *F*(1, 22) = 2.41, *p* = 0.08, *r* = .09 (see Figure 8.13). There was also no significant interaction, *F*(1.73, 38.20) < 1, $p > 0.05$, *r* = .01).

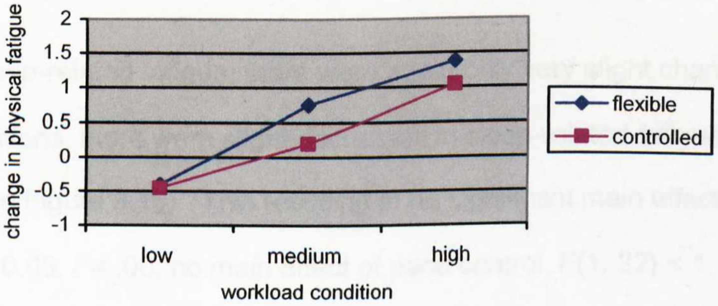


Figure 8.13: Mean change in physical fatigue as a function of workload and pace control

Changes in subjective mental fatigue:

As illustrated in Figure 8.14, there were only very small average changes in mental fatigue: In the low and medium workload conditions, participants reported slightly less mental fatigue, and slightly more under the high physical workload condition. But, despite these very small changes, the main effect of workload narrowly missed the 5% cut-off for significance *F*(2, 42) = 2.79, *p* = 0.07, *r* = .12. However, there was no main effect of pace control, *F*(1, 21) <1, $p > 0.05$, *r* = .00, and no significant interaction, *F*(2, 42) < 1, $p > 0.05$, *r* = .03.

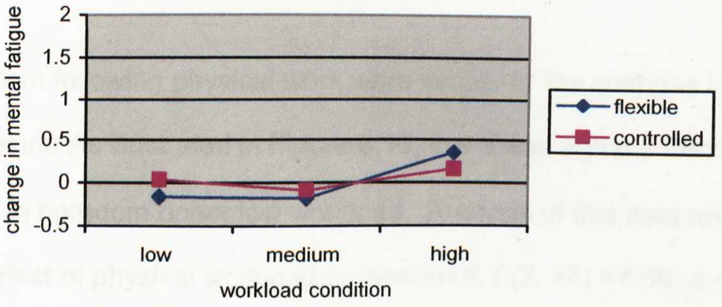


Figure 8.14: Mean change in mental fatigue as a function of physical workload and pace control

Changes in subjective sleep-related fatigue:

With regard to sleep-related fatigue, there were again only very slight changes in reported state: In all conditions, there were slight decreases in sleep-related fatigue following physical work (see Figure 8.15). This resulting in no significant main effect of workload, $F(2, 44) < 1$, $p > 0.05$, $r = .00$, no main effect of pace control, $F(1, 22) < 1$, $p > 0.05$, $r = .00$, and no significant interaction, $F(2, 44) < 1$, $p > 0.05$, $r = .00$.

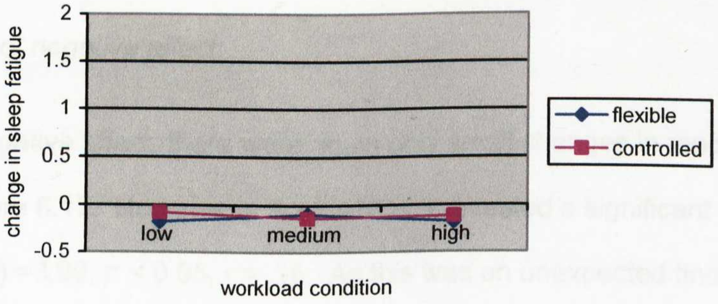


Figure 8.15: Mean change in sleep-related fatigue as a function of physical workload and pace control

Changes in level of boredom:

Changes in boredom following physical work were similar to the changes in boredom following mental work: As illustrated in Figure 8.16, there was again a trend towards greater increases in boredom under low workload. Analysis of this data revealed a significant main effect of physical workload on boredom, $F(2, 44) = 7.50$, $p < 0.01$, $r = .25$. However, there was no significant main effect of pace control, $F(1, 22) < 1$, $p > 0.05$, $r = .00$, and no significant interaction, $F(2, 44) < 1$, $p > 0.05$, $r = .00$.

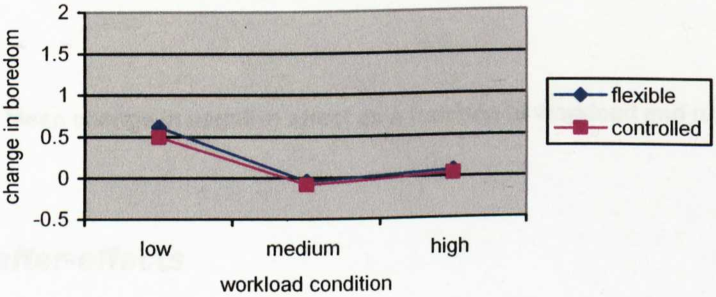


Figure 8.16: Mean change in boredom as a function of physical workload and pace control

Changes in level of negative affect:

With regard to negative affect, there were again only small changes in reported state, as illustrated by Figure 8.17. However, a 3 x 2 ANOVA revealed a significant main effect of workload, $F(2, 44) = 3.99$, $p < 0.05$, $r = .15$. As this was an unexpected finding, the data was submitted to further analysis. Post hoc Bonferroni comparisons were calculated, which revealed significant differences between only medium and high levels of workload (mean difference = -0.45, $p = 0.03$). This suggests that there is a significant difference between the change in negative affect at moderate and high levels of physical workload. Negative affect was found to reduce most under medium levels, while increasing slightly under high levels of physical workload. All other comparisons were found to be not significant ($p > 0.05$), and there was no significant main effect of pace control, $F(1,22) = 2.61$, $p > 0.05$, $r = .10$, and no significant interaction, $F(2, 44) < 1$, $p > 0.05$, $r = .02$.

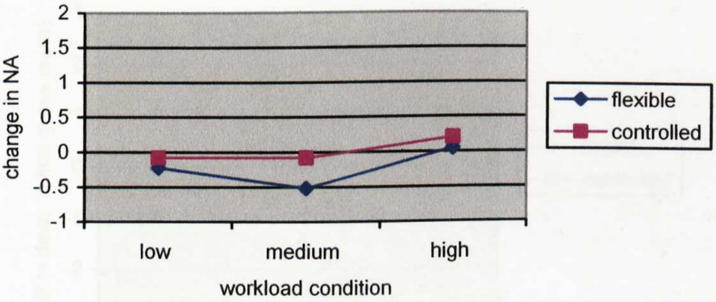


Figure 8.17: Mean change in negative affect as a function of workload and pace control

8.4.6 Fatigue after-effects

8.4.6.1 Delay until first guess

Mean delay times until the first guess were calculated for six sets of 25 fault finding trials (pre- and post-cycling; at three levels of workload). Difference scores were generated by subtracting log transformed pre- from post-cycling times and these variables were then subjected to a 3 (workload) x 2 (pace control) mixed ANOVA. As illustrated in Figure 8.18, there is a general trend towards shorter delays following physical work. While this would be expected, to some degree, in terms of a learning effect, the presence of any learning effect should be equal across groups: Therefore, if there were no effects of either IV, one would anticipate slightly negative scores equally across all conditions. However, there was a similar trend following physical work, as that identified following mental work in Chapter Seven: Participants were quicker to make their first guess following higher levels of physical work. However, these mean differences represent only small variations in actual performance, and thus the main effect of physical load on delay till first guess failed to reach significance, $F(2, 36) < 1$, $p > 0.05$, $r = .03$. There was also no main effect of pace control, $F(1, 18) < 1$, $p > 0.05$, $r = .02$., and no significant interaction, $F(2, 36) = 2.69$, $p > 0.05$, $r = .13$.

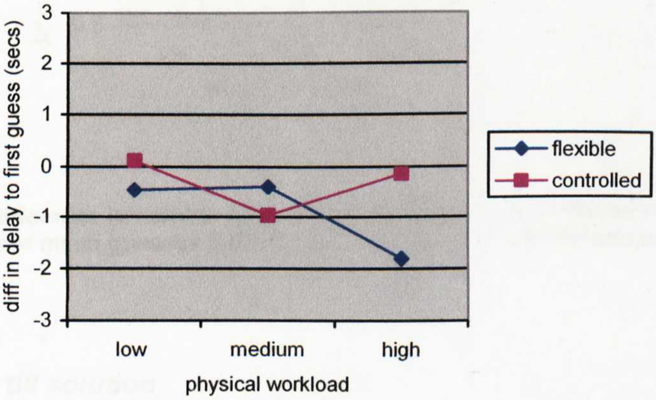


Figure 8.18: Mean difference in delay till first guess as a function of workload and pace control (a negative score indicates shorter mean delays following the loading task than before: untransformed data)

8.4.6.2 Mean number of guesses

Mean number of guesses were again calculated for six sets of 25 fault finding trials (pre- and post-cycling; at three levels of workload) and difference scores were calculated by subtracting log transformed pre- from post-cycling times. As illustrated by Figure 8.19, there are only very small changes in mean number of guesses following physical work, with a very slight trend towards a higher number of guesses for the controlled pacing group and slightly fewer guesses as a function of increasing workload. However, as one would expect, a 3 x 2 mixed ANOVA revealed no significant main effect of physical work, $F(2, 36) = 1.53$, $p > 0.05$, $r = .07$, no significant main effect of pace control, $F(1, 18) = 1.98$, $p > 0.05$, $r = .09$, and no significant interaction, $F(2, 36) < 1$, $p > 0.05$, $r = .00$.

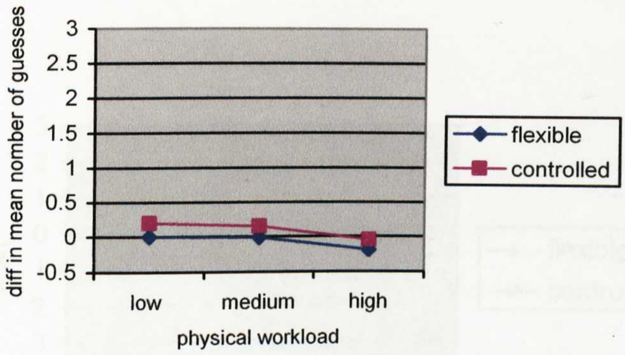


Figure 8.19: Mean difference in number of guesses as a function of workload and effort (a negative score indicates fewer mean guesses following the loading task than before: untransformed data)

8.4.6.3 Mean time till solution

Mean time till solution was calculated for six sets of 25 fault finding trials (pre- and post-cycling; at three levels of workload) and difference scores were again generated by subtracting log transformed pre- from post-cycling times. As illustrated in Figure 8.20, participants were actually faster to find the solution following cycling under higher levels of workload, and a 3 x 2 mixed ANOVA revealed there was a significant main effect of workload $F(2, 36) = 4.29$, $p > 0.05$, $r = .19$.

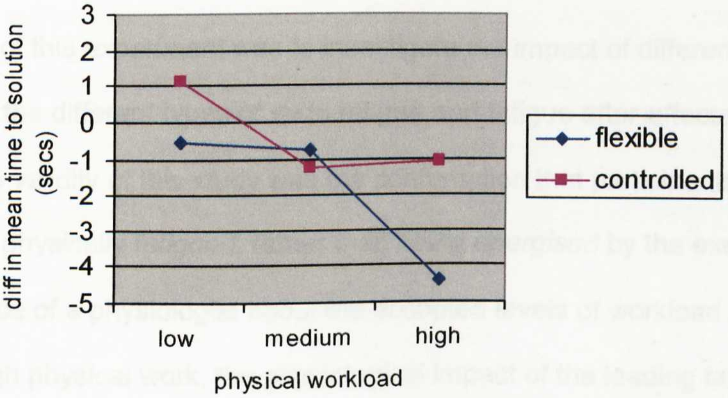


Figure 8.20: Mean difference in time till solution (post-pre) as a function of workload and effort (a negative score indicates a quicker solution following the loading task than before: untransformed data)

Figure 8.20 also illustrates a trend towards shorter time to solution following higher workload, for the flexible condition, although this effect failed to reach significance at the 5% level, $F(1, 18) = 3.20$, $p = 0.08$, $r = .15$. There was no significant interaction between workload and pace control, $F(2, 36) = 1.23$, $p > 0.05$, $r = .06$.

8.5 Discussion

The primary aim of this experiment was to investigate the impact of different levels of physical work on the different types of state fatigue and fatigue after-effects. Of primary importance to the validity of this study was the confirmation that participants could be considered to be physically *fatigued*, rather than being *energised* by the exercise. Having followed the advice of a physiologist about the accepted levels of workload relating to low, moderate and high physical work, the physiological impact of the loading task was assessed to verify that participants were working at significantly different physiological levels. In line with the multilevel analysis approach adopted throughout this thesis, this was confirmed by both the changes in subjective state and task evaluation. The overall pattern of results suggested that, in general, participants reported higher levels of effort and physical and mental demands with increasing levels of physical work. Surprisingly, while all indices of mental workload increased with the level of physical workload, there were no effects of pace control. This was despite the increased level of monitoring required to maintain a consistent pace. Furthermore, there was no effect of pace control on performance of the auditory reaction time task, which one might have expected if there were different levels of attentional resources associated with the different conditions. With the available data, it is impossible to distinguish between two potential scenarios: (1) participants in the controlled condition did in fact execute more monitoring behaviour, but this had little impact on subjective assessment and secondary task performance, or (2) there was no difference between the monitoring behaviour of the two groups, despite the differing instructions.

While this remains unclear, what is clear from the data analysis is that the experiment has been successful in manipulating significantly different levels of physical load and that the participants could be described as physically fatigued following the high level of physical load, rather than being 'aroused', 'energised' or 'activated'. Interestingly, despite the changes in physiological and subjective state, there were almost no overall effects of the physical work on subjective *mental* fatigue and *sleep-related* fatigue. Consistent with the findings relating to mental fatigue, this not only provides support for the hypothesis that individuals are capable of differentiating between different types of subjective tiredness, but consequently suggests that these different types of fatigue are differentially generated by different experiences and conditions.

What remains now is the effect of this state on the after-effects task: With regard to the delay until the first guess, participants were generally quicker to make their first guess following higher levels of workload. (While this trend failed to reach significance, this experiment possessed a fairly low level of power as a result of a small initial sample size, and some technical difficulties, which resulted in some lost data.) However, this finding was similar to the results identified for this variable in Chapter 7. But, interestingly, while the overall pattern of results in Chapter 7 suggested a shift towards low-effort processing following high levels of mental work, this pattern of results was not echoed here. There was no effect of either independent variable on number of guesses, but there was a significant effect of physical workload on time to solution: Following high physical load, participants were on average significantly faster to find the solution. This does not support a hypothesis of a speed-accuracy trade off, or a shift towards low effort processes as put forward to explain the effects following mental work. Instead, this suggests that physically fatigued participants were simply faster: This was particularly true for the flexible condition under high workload, although there was again no significant difference between the two pace control groups. Thus, the positive impact of the high level of physical workload on the after-effects task was greater for the participants with the lower level of monitoring load. This may point to positive after-effects of physical work (at this level) which are moderated by the negative after-effects of mental work. However, it is important to remember that these findings were largely general trends, rather than significant differences.

Alternatively, it is possible that, consistent with Fery *et al*'s (1997) hypothesis, the negative effects of the physical fatigue on the decision stage of the process are masked by peripheral changes. Although, this is less likely to have an impact with the current task as the speed of reaction represents only a small proportion of the delay to first guess, etc. Because the Fault Finding task is fairly complex, the average delay time is most affected by the cognitive aspects of the task.

With regard to the central question which runs throughout this thesis, these findings have some implications relating to the multidimensionality of the fatigue construct. Firstly, different conditions appear to give rise to subjectively distinct states and, secondly, these states appear to have different (specific) after-effects. This suggests some degree of different underlying processes, which is not surprising, as there has long been agreement that physical and mental fatigue differ on some level (see Hockey & Meijman, 1998). However, the extent to which these processes interact is not yet known. Therefore, the final experiment in this series will combine mental and physical work concurrently to investigate their interaction.

Chapter 9:

The interaction of physical and mental fatigue

9.1 Summary

The experiment described and discussed in the current chapter aimed to investigate the interaction of mental and physical fatigue. Mental and physical workload were each manipulated on two levels to investigate the state of fatigue which was generated by these conditions.

Again, the emphasis was on multiple levels of analysis and performance-related, subjective and physiological indices were considered to confirm that participants were 'fatigued'. In general, the pattern of subjective and objective findings supported the view that the two types of workload were sufficient to be considered as fatiguing.

The subjective findings again supported the notion that the different types of workload gave rise to subjectively distinct states which participants differentiated between. Of particular interest within the current experiment is the nature of any interaction of mental and physical workload on subjective ratings. However, no significant interactions were found. With regard to performance after-effects, the pattern of results was more consistent with the after-effects of only physical load: There were trends towards shorter delays till first guess following high workload, but this did not appear to be at a cost to time to solution.

9.2 Introduction

This final experiment represents a culmination of the work carried out so far. Psychometric and experimental methods have been employed to address the central conceptual issue of the dimensionality of the fatigue construct. What remains is an investigation of the interaction between different sources of fatigue: Essentially, the question centres around whether mental and physical work affect the same or separate mechanisms.

As discussed in Chapter 8, there is a huge body of literature investigating the effects of physical work on various psychological factors, including mental performance. These studies have differed enormously in their operationalisation of the different variables (Tomporowski & Ellis, 1986): Mental tasks have varied in nature (e.g. reaction time, decision-making, mental arithmetic) and in the timing of their administration - while some studies have measured mental performance during exertion (e.g. Stauffacher, 1937; Weingarten & Alexander, 1970), some have focused on after-effects of exercise (e.g. Davey, 1973; Gupta *et al*, 1974; Tomporowski *et al*, 1987) and others have, more comprehensively, measured mental performance both during and after exercise (e.g. McGlynn *et al*, 1977; McGlynn *et al*, 1979; Sjoberg, 1980).

The current experiment is clearly concerned with performance effects both during and after exercise. However, it differs from the earlier work (discussed in Chapter Eight) in its conceptualisation of the physical and mental work. This body of literature has traditionally addressed the effects of *exercise* on cognition. Conversely, the current series of experiments is specifically interested in the construct of *fatigue*. Therefore, rather than focusing on the exercise per se, physical work is interesting only in so far as it can be argued to generate a state of fatigue. Similarly, concurrent mental work is utilised here as a loading task, rather than an outcome variable, as used elsewhere. Therefore, in the current experiment, concurrent physical and mental work are utilised as loading tasks to generate a state of fatigue, and subjective state and performance after-effects are the dependent variables of primary interest. While these differences represent only small distinctions, it is considered important to clarify the level of relevant variables. This is particularly true considering the current conceptual framework and multiple levels of analysis because, as previously mentioned, performance effects on the CAMS task are argued to represent task engagement or disengagement, rather than fatigue effects.

Thus, while there is a body of literature in which participants have been required to carry out concurrent mental and physical tasks - this has typically conceptualised the key variables very differently, and therefore, offers little in the way of any consistent findings to inform the current research question.

9.2.1 Experiment Four rationale

Therefore, this experiment aims to consider further the issue of the multidimensionality of the fatigue construct by combining the mental and physical loading tasks (from Experiments One to Three) to investigate the nature of any effects or interactions.

To correspond with the theoretical framework of the Compensatory Control model, three levels of analysis will be addressed: (1) performance on mental loading and after-effects tasks, (2) subjective state, and (3) physiological state.

9.3 Method

9.3.1 Design

A repeated measures design was used with two independent variables: (1) Mental workload and (2) Physical workload. Mental workload was varied at two levels; *low* and *high*, determined by number of faults occurring within the CAMS computer task. Physical workload was also varied at two levels; *low* and *high*, determined by the level of resistance programmed into the stationary cycle. Both IVs were manipulated within subjects, resulting in four experimental sessions: (1) low mental / low physical load; (2) low mental / high physical load; (3) high mental / low physical load and (4) high mental / high physical load. There were twenty-four possible ways in which these four sessions could be presented and the order of session presentation was fully counterbalanced to minimise any order effects.

Each experimental session lasted approximately ninety minutes: Participants were required to complete 25 Fault Finding networks before and after one hour of CAMS management and concurrent stationary cycling.

To minimise any carryover effects, a minimum delay of three days between experimental sessions was enforced. The maximum delay was eight days.

9.3.2 Participants

The recruitment of participants for this study followed the same procedure as for Experiment 3: Participants were recruited through advertisements placed in the Sports Centre and the Graduate Research Institute at the University of Hull. Twenty students initially responded to the advert, which invited students with a good level of general fitness to take part in an experiment described as 'investigating the effects of physical exercise on performance'. Volunteers were again required to undergo an assessment of their maximum oxygen uptake (VO_2 max) carried out by a physiologist (see section 8.3.3.1). The data from this process was again used to ascertain suitability for the experiment and to provide a measure of fitness used to determine individual levels of workload. Following the fitness testing session, one participant decided not to continue with the experiment, and six further participants failed to complete the experiment; five of whom failed to meet the minimum criteria for operating the CAMS task and one who attended one of the experimental sessions but failed to attend any further sessions.

The final experimental sample of 13 consisted of 10 men and 3 women, aged 21-35 (mean = 24.98, SD = 4.52).

9.3.3 Experimental Procedure

9.3.3.1 Pre-experimental fitness testing

As stated above, participants were again required to undergo VO_2 max testing, which followed the same protocol detailed in section 8.3.3.1. This again aimed to provide a minimum fitness cut-off and allowed the calculation of individual workload, based on the objective index of fitness level.

9.3.3.2 Pre-experimental training

As with Experiment Two, this experiment required the operation of both the CAMS task and the Fault Finding Task. As all participants were new to both computer systems, a minimum of two comprehensive training sessions was required, which resulted in around four hours total training time. The training principles adopted for previous experiments were maintained, i.e. participants were again trained in small groups, with sole access to a PC.

1. *Training session 1*: Participants were initially provided with basic information about the aims and procedures of the current experiment and were then introduced to the CAMS simulator. This included a verbal description of the essential features and functions of CAMS, a cover story explaining the nature of the mission, an overview of the primary and secondary tasks and hands-on experience of manipulating the automatic and manual controls.

   Participants were then introduced to the Fault Finding task including a description of the basic features and operating principles, an explanation of possible strategies (outlined in section 5.6.3) and completion of 25 networks employing all possible strategies.

2. *Training session 2*: This involved a recap of the CAMS hierarchy of tasks and a mock experimental session of 35 minutes of system management, during which they where faced with all five system failures occurring at various intervals. Performance on this task was assessed and a maximum criterion of 2500 total time periods spent in the red and yellow zones was set. Participants who failed to meet this were required to attend a third training session which followed the procedure of training session 2. Failure to meet the minimum performance criteria resulted in the loss of five participants, all of whom spoke English as a second language, resulting in difficulty understanding the complex instructions.

   Participants were then reintroduced to the operating principles of the Fault Finding task and required to complete a further 75 networks. Performance on this task was again monitored to ensure full understanding of the task and to minimise the possibility of learning within the experimental sessions. Observation of this process revealed some difficulties in understanding the network rules. However, these were the participants who also failed to meet the criteria for the CAMS task.

9.3.3.3 Experimental Procedure

As participants were required to carry out the computer based tasks while cycling, the desktop computer was again placed on a computer stand, designed to fit around the front of the stationary cycle. This positioned the computer monitor on a platform set approximately at eye level and incorporated a lower platfrom to accommodate the keyboard and mouse and space to complete paper and pencil questionnaires.

Before starting the session, the participants were fitted with a heart rate monitor, including a chest strap and 'watch', and they were invited to adjust the seating height and handlebars of the cycle to a comfortable position. Once participants were seated on the cycle, the broad aims of the experiment and the procedure were outlined. Participants were then asked to relax for a period of two minutes, in order to provide a baseline measure of resting heart rate.

The participants then completed 25 Fault Finding networks, followed by the state fatigue questionnaire (FQ). For the following hour they were required to complete three tasks: Continuous management of CAMS; 3 x 15 minute periods of cycling; and verbal response to mid-task questionnaires. (See Table 9.1).

This one-hour period of mental and physical workload was followed by the completion of the paper and pencil state questionnaires and, finally, 25 additional Fault Finding networks.

Time Period	Minutes	*CAMS*	*Cycling*	*Mid-task questionnaires*
1	0-5	Start		
	5-20	Faults occur	▨	18 minutes
2	20-25	All faults in place	Rest	
	25-40	All faults in place	▨	38 minutes
3	40-45	All faults in place	Rest	
	45-60	All faults in place	▨	58 minutes

Table 9.1: Experimental session - Concurrent activities

9.3.4 Independent Variables

This study involved the manipulation of two repeated measures IVs: mental and physical workload.

Mental Workload was varied at two levels: *low* and *high*. The number of system failures selected to represent each of these levels of workload was again based on the findings of Experiment 1; two system failures were selected to represent *low* workload, and all five system failures represented *high* workload. However, following from the findings of Experiment 2, all participants were given the 'high effort' instructions and urged to 'invest the maximum effort into the CAMS tasks, as this was vital to the success of the experiment'. As participants would experience each level of mental load twice, two equivalent experimental schedules were designed for each level of workload to minimise practice effects (see Table 9.2). Presentation of schedule A or B was randomised.

	Fault presentation			
Workload condition	***Schedule A***	(mins)	***Schedule B***	(mins)
Low workload	Oxygen	*3*	CO^2	*3*
	Temperature	*20*	Humidity	*20*
High workload	CO^2	*3*	Oxygen	*3*
	Oxygen	*7*	CO^2	*7*
	Temperature	*13*	Humidity	*13*
	Humidity	*16*	Temperature	*16*
	Pressure	*20*	Pressure	*20*

Table 9.2: Experimental schedule: Fault timetables

Across both levels of workload, the first fault was presented at three minutes and final fault was presented at twenty minutes. Therefore, between twenty and sixty minutes, all occurring faults were present.

Physical workload was varied at two levels: *low* and *high*, manipulated by varying the intensity of the physical demands. As with Experiment 3, this was achieved by maintaining a speed of 60 RPM at both workload levels, while adjusting the resistance by adding the appropriate weight to the cycle. The *high* workload condition required participants to work for 3 x 15 minute periods at 70% of their maximum capacity, and *low* workload required the participants to work for 3 x 15 minute periods of 30% of VO_2 max. In some cases this was less than the minimum adjustable level. In such cases, participants were required to maintain a speed of 60 RPM with the minimum weight resistance (only the weight of the basket). The breaks in the cycling were provided following the findings of a simple pilot study of the experimental session. An individual of comparable fitness to the participants was required to cycle for the full hour. This participant was not able to complete the hour of continuous cycling at the high physical workload level, thus it was considered appropriate to provide two short breaks of five minutes duration. Participants were instructed when to start and stop cycling.

9.3.5 Dependent Variables

9.3.5.1 Performance-related dependent variables

CAMS Task

The CAMS task again provided the broad range of dependent variables detailed in Chapter Six (see section 6.3.5.1). In summary, this includes performance on the primary task (minor and major parameter control failures) and performance on the secondary tasks (delay in alarm acknowledgement and timing of system checks).

Fault Finding task

Dependent variables provided by the Fault Finding task include measures of (1) mean time until first guess (2) mean time until solution and (3) mean number of guesses.

9.3.5.2 Subjective measures

As with the previous experiments, state fatigue was measured with the 15 item *Feelings Questionnaire* and the *Workload Questionnaire* was used to evaluate the nature of the demands associated with the task. This experiment also incorporated the brief mid-task evaluation questionnaire (outlined in section 8.3.5.2) which provided indices of physical demands, mental demands, tiredness and effort. This scale was presented at 3 intervals during the cycling: after 18 minutes, 38 minutes and 58 minutes. Participants were again given the option of responding verbally or by pointing at the number which best reflected their rating.

9.3.5.3 Physiological measures

Heart rate was measured throughout the entire experimental session via a pulse monitor, which required participants to wear a chest strap, and a receiver device worn on the wrist.

9.3.6 Data treatment

As with the previous experiments, Analysis of Variance was the main method of data analysis and performance data from both computer tasks were corrected using (1) log transformation, and (2) adjustment of the degrees of freedom, where sphericity had been violated.

With regard to the Fault Finding task, difference variables were calculated, by subtracting pre- from post-cycling averages.

The questionnaire data were again reduced into their relevant factors: (1) A single index of overall workload, and (2) five factors of subjective state: mental, physical and sleep-related fatigue, boredom and negative affect. Difference scores were then obtained by subtracting pre- from post-factor scores.

Cardiac activity data was again broken down into 3 time periods (1) during the Fault Finding task – prior to the loading tasks, (2) during the loading tasks, and (3) during the Fault Finding task – following the cycling.

9.4 Results

The primary aim of this experiment was to investigate the development of fatigue following two levels of mental and two levels of physical workload. Within the context of the theoretical framework, a multilevel analysis is required. This will involve: performance on the mental loading task; physiological data; changes in subjective state and task evaluation; and performance after-effects.

9.4.1 Mental Loading Task: CAMS Performance

9.4.1.1 Primary task performance: Parameter Control Failures

Performance on the primary task was again scored in terms of percentage of time each parameter was out of range. As performance on this task is, by nature, so closely related to number of system failures, it was expected that participants would have fewer PCFs under both conditions involving low mental workload, irrespective of physical workload and performance protection.

As expected, Figure 9.1 illustrates a pattern of results where primary task performance under the two low mental workload conditions was superior to performance under the two conditions involving high mental workload. Interestingly, there appears to be little effect of physical workload on primary task performance, although there is a slight trend for performance under low physical workload to improve in period three and deteriorate in period three under high physical load. However, these differences are very small.[1] Overall, though, performance could again be argued to have been largely protected, as average parameter control performance remained within the normal range for more than 90% of the time, even in the high mental workload conditions which involved the manual control of all five faults.

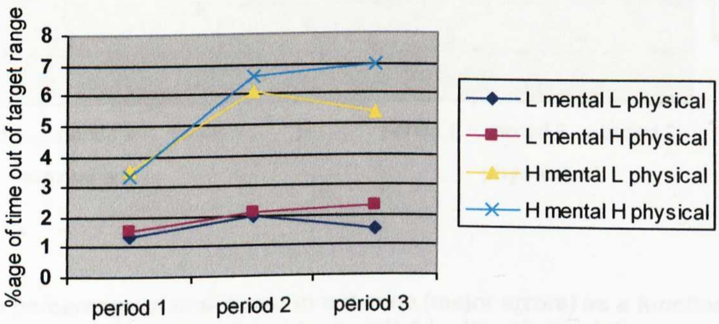


Figure 9.1: Mean percentage of time spent out of range (major + minor errors)

To investigate primary task performance further, PCFs were classified as either major errors (time spent in the red zone) or minor errors (time spent in yellow zone) and subjected to further analysis.

A 2 (mental workload) x 2 (physical workload) x 3 (time period) mixed design ANOVA was first carried out on the log transformed major errors. As expected, this revealed a highly significant main effect of mental workload $F(1, 12) = 64.89$, $p < 0.001$, $r = .84$, with those in high mental workload conditions committing a higher number of major PCFs (see

---

[1] In interpreting Figure 9.1 and other CAMS performance data, it is important to note that, in all four conditions, all programmed faults were in place at the end of period one. Therefore, any differences between periods one and two could be a product of different levels of workload, while there is no difference between the workload in periods two and three.

Figure 9.2). However, there were no significant main effects of physical workload, $F(1, 12) < 1$, $p > 0.05$, $r = .03$, or time period, $F(2, 24) = 2.36$, $p > 0.05$, $r = .16$. Although, Figure 9.2 does show a trend towards a higher number of major errors under high physical load.

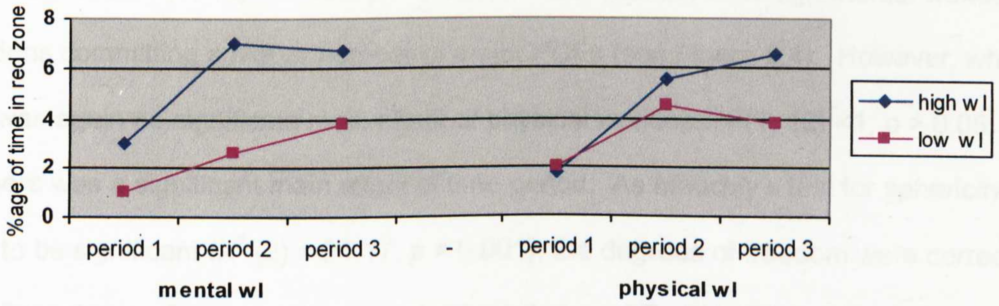


Figure 9.2: Mean percentage of time spent in red zone (major errors) as a function of mental and physical workload and time period (untransformed data)

There was also a significant interaction of mental and physical workload, $F(1, 12) = 4.93$, $p < 0.05$, $r = .29$ (see Figure 9.3), with the negative effect of physical load being greater at high levels of mental load. However, there were no further interactions: mental x period, $F(2, 24) < 1$, $p > 0.05$, $r = .04$; physical x period, $F(2, 24) < 1$, $p > 0.05$, $r = .00$; mental x physical x period, $F(2, 24) = 2.02$, $p > 0.05$, $r = .14$

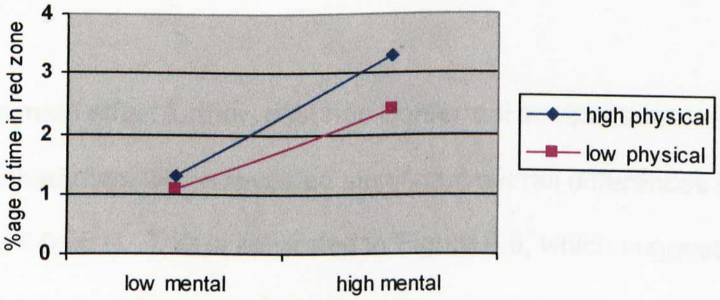


Figure 9.3: Mean percentage of time spent in red zone (major errors) as a function of mental and physical workload (untransformed data)

A slightly different pattern of results was found having subjected minor parameter control failures to a 2 x 2 x 3 mixed ANOVA. There was again a highly significant main effect of mental workload, $F(1, 12) = 71.08$, $p < 0.001$, $r = .85$, with those in high mental workload conditions committing a higher number of major PCFs (see Figure 9.4). However, while there was again no significant main effect of physical workload, $F(1, 12) < 1$, $p > 0.05$, $r = .06$, there was a significant main effect of time period: As Mauchly's test for sphericity was found to be significant (X^2 (2) = 21.17, $p < 0.001$), the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .55$), $F(1.10, 13.21) = 6.26$, $p < 0.05$, $r = .34$.

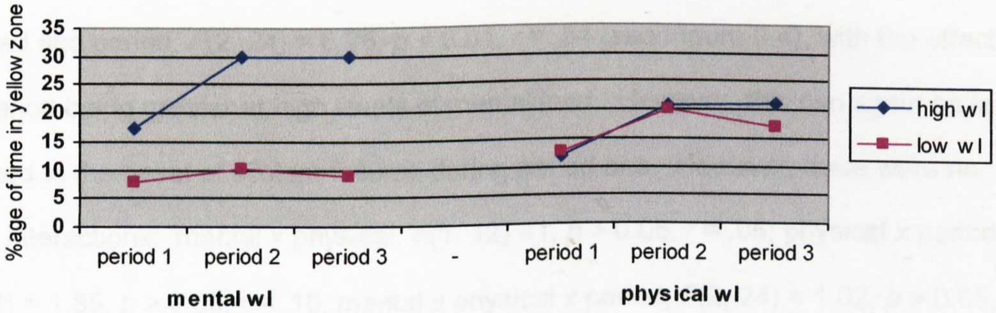


Figure 9.4: Mean percentage of time spent in yellow zone (minor errors) as a function of mental and physical workload and time period (untransformed data)

To investigate this main effect further, post hoc Bonferroni comparisons were calculated on the log transformed data, which revealed significant overall differences only between periods 1 and 2 ($p < 0.001$). This is illustrated in Figure 9.5, which suggests a rise in errors from period 1 to 2, which can largely be attributed to the presentation of faults during period 1.

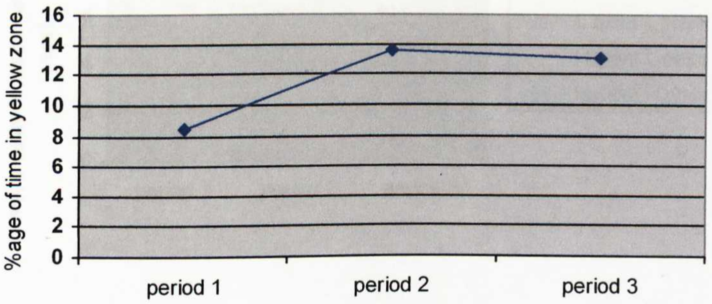


Figure 9.5: Mean overall percentage of time spent in yellow zone (minor errors) as a function of time period (untransformed data)

In addition to these main effects, there was also a significant interaction of mental workload and period, *F*(2, 24) = 6.26, $p < 0.01$, *r* = .34 (see Figure 9.4), with the effect of time period being greater at high levels of mental load. However, this can again be largely attributed to the onset of system failures during period one. However, there were no further interactions: mental x physical, *F*(1, 12) <1, $p > 0.05$, *r* = .05; physical x period, *F*(2, 24) = 1.35, $p > 0.05$, *r* = .10; mental x physical x period, *F*(2, 24) = 1.02, $p > 0.05$, *r* = .07.

9.4.1.2 Secondary task performance

Alarm response time

Although one would expect an impact of workload on this secondary task, results from previous experiments have not always been as expected. As illustrated in Figure 9.6, the condition of low mental and low physical workload yielded fairly low response time, but there was little pattern in the average response times for the remaining conditions.

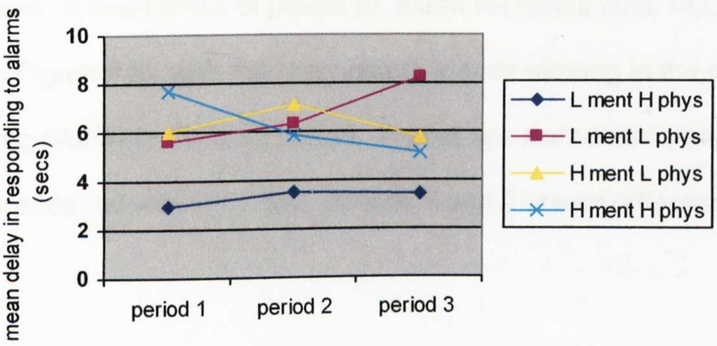


Figure 9.6: Mean alarm response times by workload condition (untransformed data)

To analyse this data further, response times, across three time periods, were log transformed and subjected to a 2 x 2 x 3 mixed design ANOVA. Although one would expect a significant main effect of mental workload, this result did not quite reach significance, $F(1,12) = 3.51$, $p = 0.08$, $r = .26$ (see Figure 9.7). The trend here was for response times to be slower, overall, under high mental workload. There was also also no main effect of physical workload, $F(1,12) < 1$, $p > 0.05$, $r = .00$, although, response times are again typically slower under high physical load.

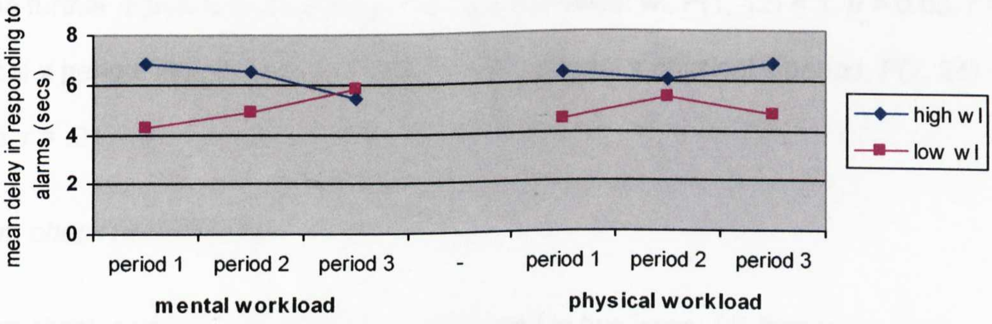


Figure 9.7: Mean alarm response time as a function of mental and physical workload and time period (untransformed data)

There was, however, a main effect of period on alarm response time, $F(2, 24) = 4.16$, $p < 0.05$, $r = .25$ (see Figure 9.8), with the response to alarms slowing in the second period, but then increasing slightly in the third period. A post hoc Bonferroni comparison revealed a significant difference between only time periods 1 and 2 (mean difference = 0.13, $p < 0.05$)

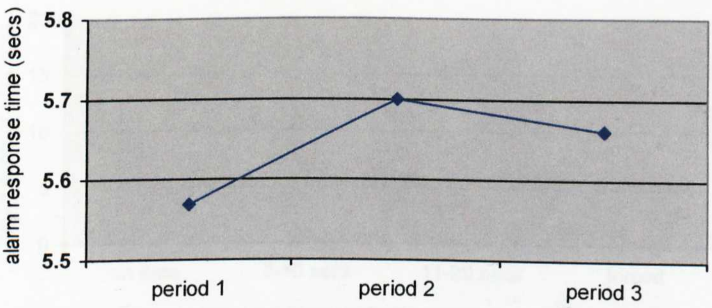


Figure 9.8: Mean alarm response time (secs) as a function of time period (untransformed data)

Although Figure 9.7 suggests a trend towards an interaction between mental workload and time period, this failed to reach significance, $F(2, 24) = 1.23$, $p > 0.05$, $r = .09$. There were also no further significant interactions: mental x physical wl, $F(1, 12) < 1$, $p > 0.05$, $r = .02$; physical x period, $F(2, 24) < 1$, $p > 0.05$, $r = .05$; mental x physical x period, $F(2, 24) < 1$, $p > 0.05$, $r = .01$.

System check performance:

System check performance again was calculated in two ways: (1) *frequency* of four categories of delay as either (i) on time (0-2 second delay), (ii) 'slightly' late (3-10 seconds delay), (iii) 'very' late (11-29 seconds delay) or (iv) forced checks (generated by the system at 30 seconds late); (2) *mean* time delay to system check.

Figure 9.9 illustrates the patterns of performance for this secondary task under the four workload conditions. This reveals a coherent pattern of an effect of mental workload: the two low mental conditions follow a consistent pattern, with a high proportion of on time checks. The two high mental workload conditions also show a consistent pattern, with fewer on time checks and a slightly higher number of later checks in each category. There appears to be little effect of physical load on system check performance.

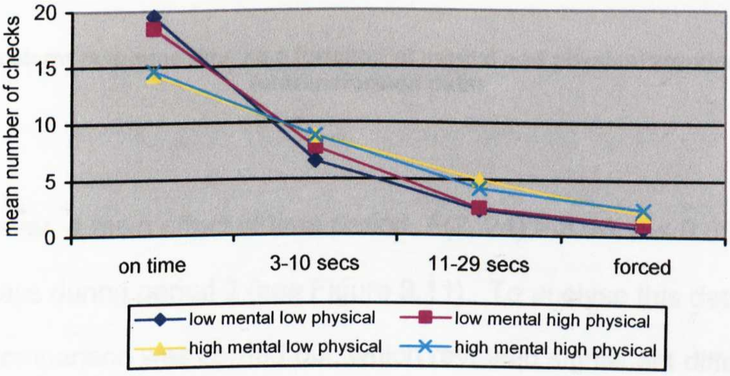


Figure 9.9: Mean number of checks in each category of system check performance

To investigate this further, mean system check delays were log transformed and subjected to a 2 x 2 x 3 mixed design ANOVA. Unlike the alarm response secondary task, this analysis did reveal a highly significant main effect of mental workload, $F(1, 12) = 23.12$, $p < 0.001$, $r = .65$, with a higher mean delay in system checks under conditions of high mental workload (see Figure 9.10). Figure 9.10 also reveals that there was no main effect of physical workload, $F(1, 12) < 1$, $p > 0.05$, $r = .01$.

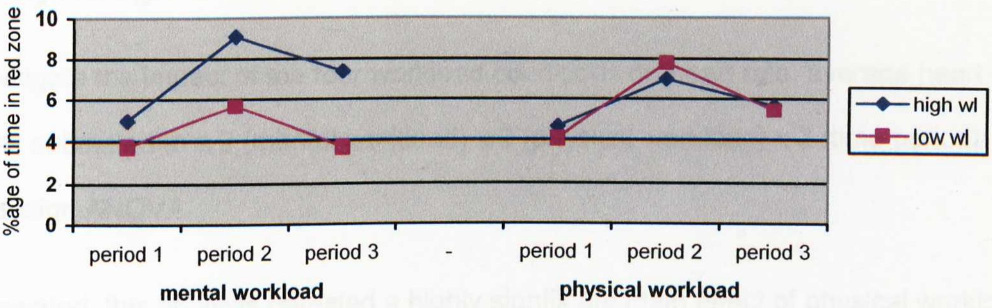


Figure 9.10: Mean alarm response time as a function of mental and physical workload and time period (untransformed data)

There was, however, a main effect of time period, $F(2, 24) = 8.30$, $p < 0.01$, $r = .40$, with longer mean delays during period 2 (see Figure 9.11). To analyse this data further, a post hoc Bonferroni comparison was carried out, which revealed significant differences between periods 1 and 2 (mean difference = 3.05, $p < 0.001$) and 2 and 3 (mean difference = 1.87, $p < 0.05$), which suggests that participants got significantly slower in the second period, when the full mental workload was in place, but then improved again in the third period. There were no significant interactions: mental workload x period, $F(2, 24) = 2.47$, $p > 0.05$, $r = .17$; mental x physical workload, $F(1, 12) < 1$, $p > 0.05$, $r = .00$; physical x period, $F(2, 24) < 1$, $p > 0.05$, $r = .03$; mental x physical x period, $F(2, 24) < 1$, $p > 0.05$, $r = .00$.

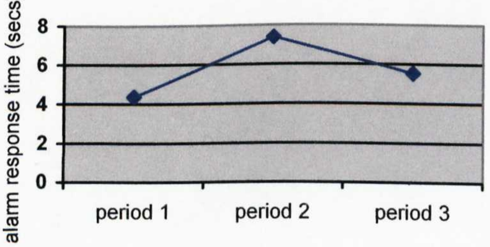


Figure 9.11: Mean delay in system checks (secs) as a function of time period (untransformed data)

9.4.2 Physiological data

To investigate the impact of the four workload conditions on heart rate, average heart rate data was submitted to a 2 (mental workload) x 2 (physical workload) x 3 (time period) mixed design ANOVA.

As anticipated, this analysis revealed a highly significant main effect of physical workload, $F(1, 12) = 108.13$, $p < 0.001$, $r = .90$, with higher average heart rates under conditions of high physical load (see Figure 9.12). This importantly confirms that participants were working significantly physiologically harder under high physical workload. There was no main effect of mental workload, $F(1, 12) < 1$, $p > 0.05$, $r = .00$, as average heart rates were no different at the two levels of mental load, for each time period. There was, however, a highly significant main effect of time period, $F(2, 24) = 238.57$, $p < 0.001$, $r = .95$, with heart rates beginning low during the first Fault Finding period (mean = 80.53, sd = 11.70), rising during the cycling in period 2 (mean = 122.78, sd = 12.96), and falling again when participants were completing the second series of Fault Finding networks in period 3 (mean = 85.46, sd = 13.26). To investigate where the significant differences lay, the heart rate data was then submitted to a Bonferroni post-hoc test. This revealed highly significant differences between periods 1 and 2 (mean difference = 42.25, $p < 0.001$) and periods 2 and 3 (mean difference = 37.32, $p < 0.001$) but not between periods 1 & 3 (mean difference = 4.91, $p > 0.05$).

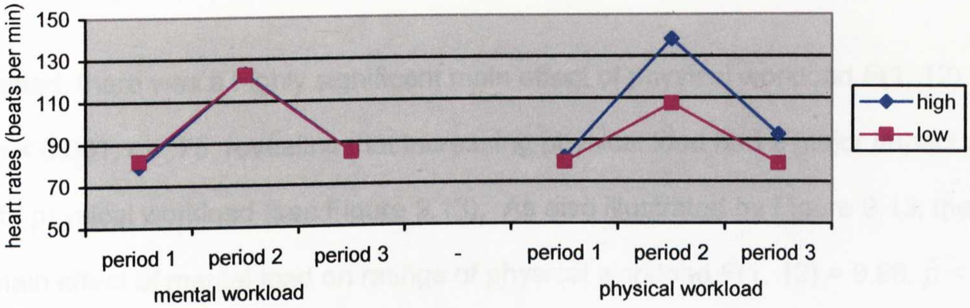


Figure 9.12: Mean heart rate as a function of mental and physical workload and time period

As also suggested by Figure 9.12, there was no significant interaction between mental workload and time period, $F(2, 24) = 1.02$, $p > 0.05$, $r = .07$. However, Figure 9.12 does illustrate the significant interaction between physical workload and period, $F(2, 24) = 114.78$, $p < 0.001$, $r = .90$, with the effect of physical workload being greatest in period 2, during the cycling. There were no further significant interactions: mental workload x physical workload, $F(1, 12) < 1$, $p > 0.05$, $r = .0$; mental workload x physical workload x period, $F(2, 24) = 2.96$, $p > 0.05$, $r = .19$.

9.4.3 Subjective task evaluation

9.4.3.1 Subjective mid-task workload assessment

To investigate subjective evaluations of the concurrent mental and physical loading task, the ratings of physical workload, mental workload, effort and tiredness were subjected to a 2 x 2 x 3 mixed design ANOVA.

Concurrent ratings of physical workload

As expected, there was a highly significant main effect of physical workload $F(1, 12) = 36.17$, $p < 0.001$, $r = .75$, revealing that increasing physical load had a major impact on ratings of physical workload (see Figure 9.13). As also illustrated by Figure 9.13, there was a main effect of mental load on ratings of physical workload $F(1, 12) = 9.98$, $p < 0.01$, $r = .45$, revealing that ratings of physical workload were higher under high mental work.

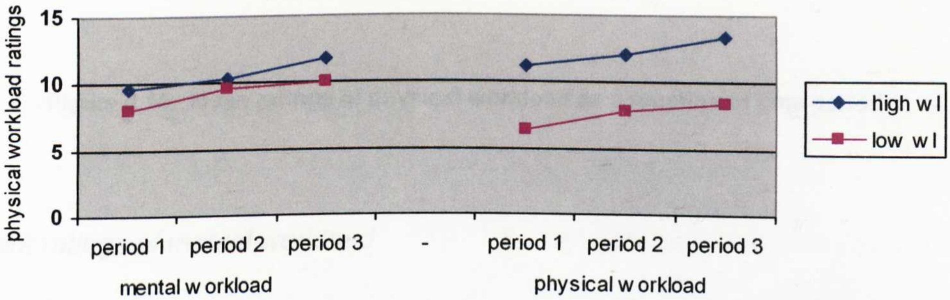


Figure 9.13: Mean physical workload ratings as a function of physical load, mental load and period

With regard to the analysis of time period, Mauchly's test of sphericity was found to be significant (X^2 (2) = 17.74, $p < 0.01$) so the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .60$). This resulted in a highly significant main effect of period on ratings of physical load, $F(1.26, 15.23) = 27.29$, $p < 0.001$, $r = .69$, with physical ratings of increasing throughout the session (see Figure 9.14), Post hoc Bonferroni comparisons revealed significant differences between all levels of period; 1 and 2 (mean difference = 1.19, $p < 0.001$); 2 and 3 (mean difference = 0.94, $p < 0.05$); 1 and 3 (mean difference = 2.13, $p < 0.001$). There were, however, no significant interactions:

mental workload x period $F(2, 24) = 3.22$, $p > 0.05$, $r = .21$; mental x physical workload, $F(1, 12) = 1.11$, $p > 0.05$, $r = .08$; physical load x period, $F(1, 12) < 1$, $p > 0.05$, $r = .06$; workload x period x condition, $F(2, 24) < 1$, $p > 0.05$, $r = .05$.

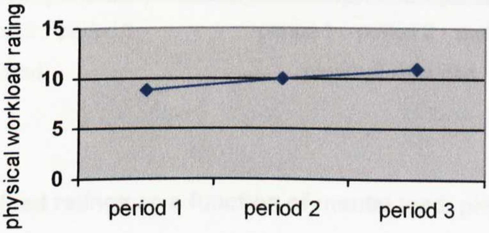


Figure 9.14: Mean ratings of physical workload as a function of time period

Concurrent ratings of mental workload

A similar pattern of results was found for the effects of mental workload on subjective task ratings. As expected, a 2 x 2 x 3 mixed ANOVA revealed a highly significant main effect of mental workload $F(1, 11) = 18.98$, $p < 0.001$, $r = .63$. As illustrated in Figure 9.15, increasing mental workload had a major impact on ratings of mental workload. Also illustrated in Figure 9.15 is the main effect of physical load on ratings of mental workload $F(1, 12) = 5.63$, $p < 0.05$, $r = .33$, which reveals that ratings of mental workload were also higher under high physical work.

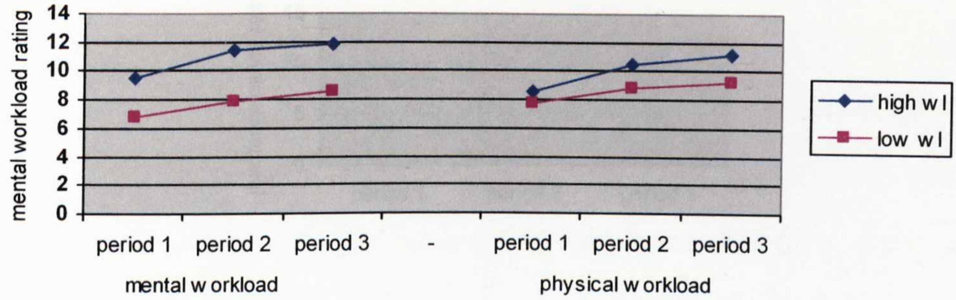


Figure 9.15: Mean mental workload ratings as a function of mental load, physical load and period

With regard to the analysis of time period, Mauchly's test of sphericity was again found to be significant (X^2 (2) = 6.59, $p < 0.05$) so the degrees of freedom were corrected using Huynh-Feldt estimates of sphericity (ε = .76). This resulted in a significant main effect of period on ratings of mental load, $F(1.35, 14.87) = 13.73$, $p < 0.001$, $r = .55$, with physical ratings increasing throughout the session (see Figure 9.16). But, while post hoc Bonferroni comparisons revealed significant differences between periods 1 and 2 (mean difference = 1.50, $p < 0.01$ and 1 and 3 (mean difference = 2.02, $p < 0.01$), the difference between 2 and 3 was not found to be significant (difference = 0.52, $p > 0.05$). These differences are consistent with the increasing demands from periods 1 to 2 and stable demands from periods 2 to 3. There were also no significant interactions: mental workload x period $F(2, 22) = 1.37$, $p > 0.05$, $r = .11$; mental x physical workload, $F(1, 11) < 1$, $p > 0.05$, $r = .06$; physical load x period, $F(2, 22) = 1.23$, $p > 0.05$, $r = .10$; workload x period x condition, $F(2, 22) = 1.86$, $p > 0.05$, $r = .14$.

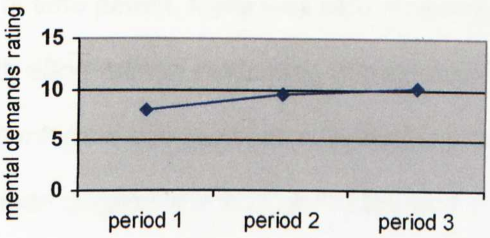


Figure 9.16: Mean ratings of mental workload as a function of time period.

Concurrent ratings of effort

Concurrent ratings of effort were also subjected to a 2 x 2 x 3 mixed ANOVA. As expected, this revealed a significant main effect of mental workload $F(1, 11) = 11.41$, $p < 0.01$, $r = .50$. As illustrated in Figure 9.17, high mental workload resulted in higher mean ratings of effort. Similarly, Figure 9.17 also illustrates the main effect of physical load on ratings of effort, $F(1, 11) = 22.66$, $p < 0.001$, $r = .67$, with ratings of effort being significantly higher under high physical work.

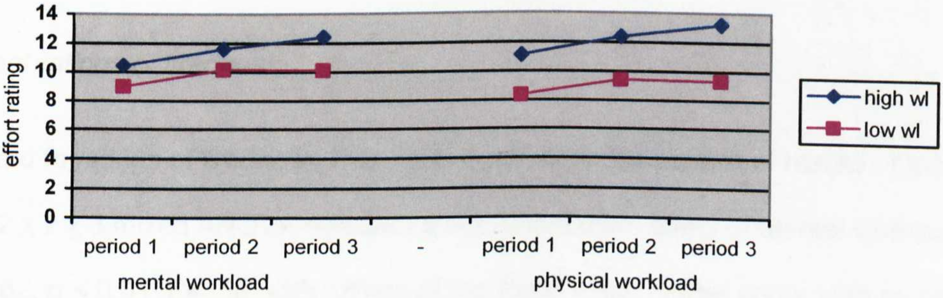


Figure 9.17: Mean effort ratings as a function of mental load, physical load and period

With regard to the analysis of time period, there was also a significant main effect, $F(2, 22) = 7.94$, $p < 0.01$, $r = .41$, with effort ratings increasing throughout the session (see Figure 9.18). But, while post hoc Bonferroni comparisons revealed significant differences between periods 1 and 2 (mean difference = 1.16, $p < 0.05$) and 1 and 3 (mean difference = 1.54, $p < 0.05$), the difference between 2 and 3 was again found to be not significant (difference = 0.37, $p > 0.05$). There were also again no significant interactions: mental workload x period $F(2, 22) = 1.64$, $p > 0.05$, $r = .13$; mental x physical workload, $F(1, 11) < 1$, $p > 0.05$, $r = .02$; physical load x period, $F(2, 22) = 1.75$, $p > 0.05$, $r = .13$; workload x period x condition, $F(2, 22) < 1$, $p > 0.05$, $r = .04$.

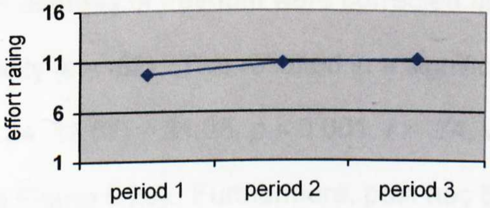


Figure 9.18: Mean ratings of mental workload as a function of time period.

Concurrent ratings of tiredness

With regard to ratings of tiredness, there was again a similar pattern of results. Once again, a 2 x 2 x 3 mixed ANOVA revealed a significant main effect of mental workload $F(1, 11) = 15.62$, $p < 0.01$, $r = .58$, with ratings of tiredness being higher under high mental workload (see Figure 9.19). Also illustrated in Figure 9.19 is the highly significant main effect of physical workload on ratings of tiredness, $F(1, 11) = 39.19$, $p < 0.001$, $r = .78$, which reveals that ratings of tiredness were also higher under high physical workload.

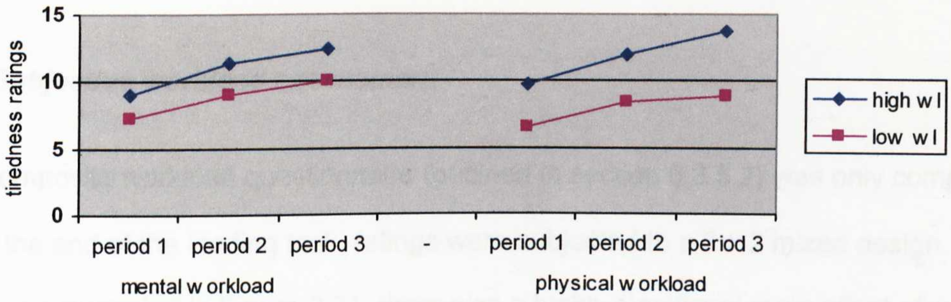


Figure 9.19: Mean tiredness ratings as a function of mental load, physical load and period

Mauchly's test of sphericity was again found to be significant for the analysis of period (X^2 (2) = 11.69, $p < 0.01$) so the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .62$). This resulted in a significant main effect of period on ratings of tiredness, $F(1.24, 13.67) = 31.26$, $p < 0.001$, $r = .74$, with tiredness increasing throughout the session (see Figure 9.20). Furthermore, post hoc Bonferroni comparisons revealed significant differences between all periods: 1 and 2 (mean difference = 2.08, $p < 0.001$); 1 and 3 (mean difference = 3.10, $p < 0.001$); and 2 and 3 (mean difference = 1.02, $p > 0.05$). However, there were also no significant interactions: mental workload x period $F(2, 22) < 1$, $p > 0.05$, $r = .04$; mental x physical workload, $F(1, 11) < 1$, $p > 0.05$, $r = .05$; physical load x period, $F(2, 22) = 2.27$, $p > 0.05$, $r = .17$; workload x period x condition, $F(2, 22) < 1$, $p > 0.05$, $r = .02$.

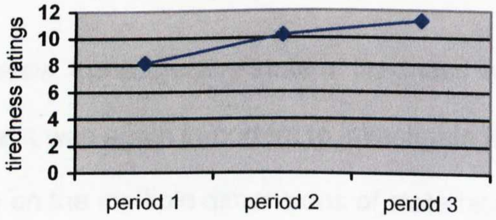


Figure 9.20: Mean ratings of mental workload as a function of time period

9.4.3.2 Subjective workload assessment

As the composite workload questionnaire (outlined in section 6.3.5.2) was only completed once, at the end of the loading task, ratings were subjected to a 2 x 2 mixed design ANOVA. As illustrated in Figure 9.21, there was a highly significant main effect of mental workload on overall workload assessment, $F(1, 11) = 29.53$, $p < 0.001$, $r = .72$, revealing ratings of overall workload were higher under high mental workload. Furthermore, there was also a highly significant main effect of physical load, $F(1, 11) = 17.07$, $p < 0.01$, $r = .60$, again with higher overall ratings of workload under conditions of high physical load. There was, however, no interaction, $F(1,11) < 1$, $p > 0.05$, $r = .05$.

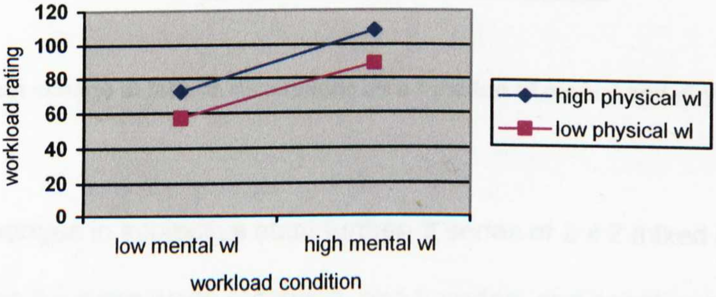


Figure 9.21: Mean workload ratings as a function of mental and physical workload

9.4.4 Subjective state

As stated throughout this thesis, the subjective state of tiredness is a central component of fatigue research. Therefore, it was again important to investigate the impact of the experimental manipulations on the multiple dimensions of state fatigue. Having calculated a difference variable for each of the fatigue states, changes in mental, physical and sleep-related fatigue were considered. As illustrated in Figure 9.22, the two workload IVs again had a differential impact on these three fatigue dimensions. There was little effect of the manipulations on sleep-related fatigue. However, there were increases in mental fatigue under high mental load and increases in physical fatigue under high physical load. The only surprising finding was that the increase in physical fatigue was greatest when mental workload was low, although this was only a small difference.

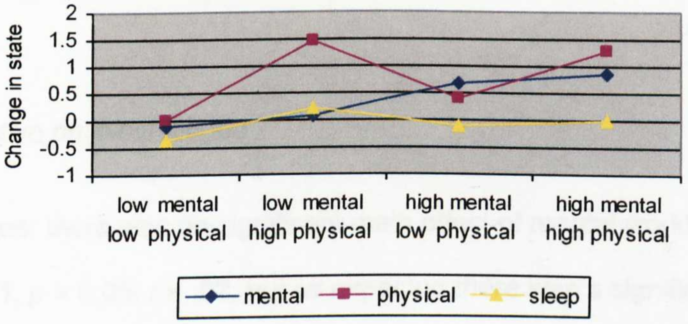


Figure 9.22: Mean change in fatigue dimensions as a function of mental and physical workload

To analyse the changes in subjective state further, a series of 2 x 2 mixed design ANOVAs were carried out on the three types of fatigue, and boredom and negative affect.

Changes in subjective mental fatigue

As expected there was a highly significant main effect of mental workload $F(1, 11) = 20.99$, $p < 0.001$, $r = .65$, revealing that ratings of mental fatigue increased by a significantly larger amount under high mental workload (see Figure 9.23). However, there was no main effect of physical workload on ratings of mental fatigue, $F(1, 11) = 1.38$, $p > 0.05$, $r = .11$, although increases in mental fatigue were higher under high physical load. There was no significant interaction, $F(1, 11) < 1$, $p > 0.05$, $r = .00$.

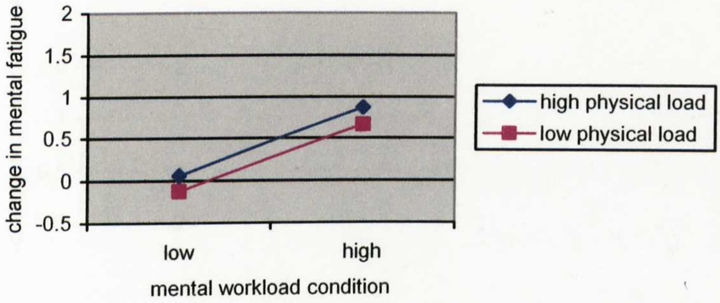


Figure 9.23: Mean change in mental fatigue as a function of mental and physical workload

Changes in subjective physical fatigue

Unlike mental fatigue, there was no significant main effect of mental workload on physical fatigue, $F(1, 11) < 1$, $p > 0.05$, $r = .02$, but as expected there was a significant main effect of physical workload on physical fatigue, $F(1, 11) = 16.64$, $p < 0.01$, $r = .60$ (see Figure 9.24). However, while the interaction between workload and effort failed to reach significance ($F(1, 11) = 2.31$, $p = 0.08$, $r = .17$), Figure 9.24 illustrates a slight trend towards an interaction: the effect of physical work is greater under low mental work. While these numbers represent fairly small changes on a 1-5 Likert scale, it is possible that the high mental / high physical difference score may actually be revealing a ceiling effect. To discount this possibility, the individual post high/high scores were investigated and it was revealed that the modal score following high/high workload was a rating of 4. Therefore, one could not explain this effect as a statistical artefact.

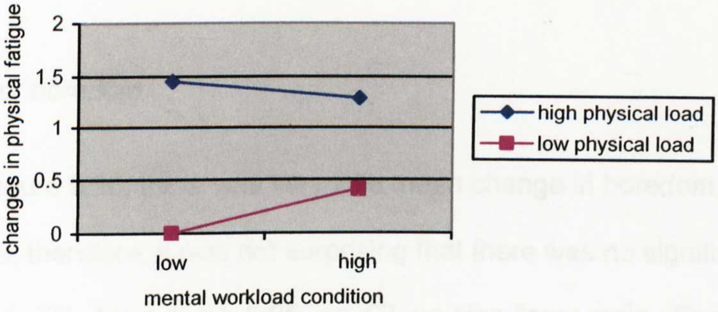


Figure 9.24: Mean change in physical fatigue as a function of mental and physical workload

Changes in subjective sleep-related fatigue

With regard to sleep-related fatigue, there was no significant main effect of mental workload, *F*(1, 11) <1, $p > 0.05$, $r = .00$. However, there was a significant main effect of physical load, *F*(1, 11) = 5.17, $p < 0.05$, $r = .32$, with an overall decrease in sleep-related fatigue following low physical load and an overall increase following high physical load (see Figure 9.25). Interestingly, Figure 9.25 reveals the same trend towards an interaction as found for physical fatigue: the impact of physical load is lower when under high mental load. But, again, these differences are very small on the 1-5 scale.

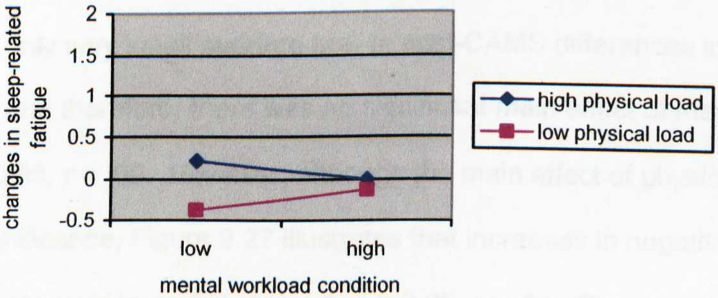


Figure 9.25: Mean change in sleep-related fatigue as a function of mental and physical workload

Changes in level of boredom

As illustrated in Figure 9.26, there was very little mean change in boredom, within any of the four conditions, therefore, it was not surprising that there was no significant main effect of mental workload, $F(1, 11) < 1$, $p > 0.05$, $r = .02$, no significant main effect of physical workload, $F(1, 11) < 1$, $p > 0.05$, $r = .00$, and no significant interaction, $F(1, 11) < 1$, $p > 0.05$, $r = .04$.

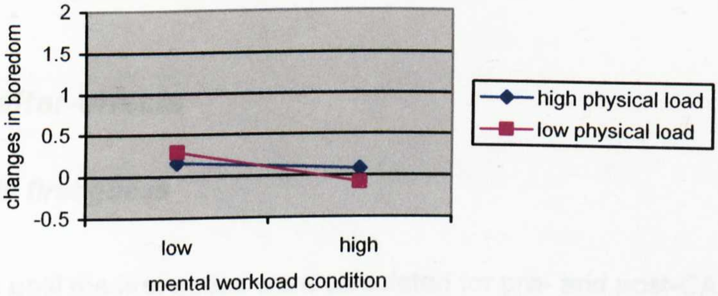


Figure 9.26: Mean change in boredom as a function of mental and physical workload

Changes in level of negative affect

Again there were only very small average pre- to post-CAMS differences in negative affect (see Figure 9.27) and therefore, there was no significant main effect of mental workload, $F(1, 11) < 1$, $p > 0.05$, $r = .00$. However, although the main effect of physical workload failed to reach significance, Figure 9.27 illustrates that increases in negative affect were higher under high physical load, $F(1, 11) < 1$, $p > 0.05$, $r = .01$. There was also a similar trend towards an interaction as identified for physical and sleep related fatigue, but this again failed to reach significance, $F(1, 11) < 1$, $p > 0.05$, $r = .07$.

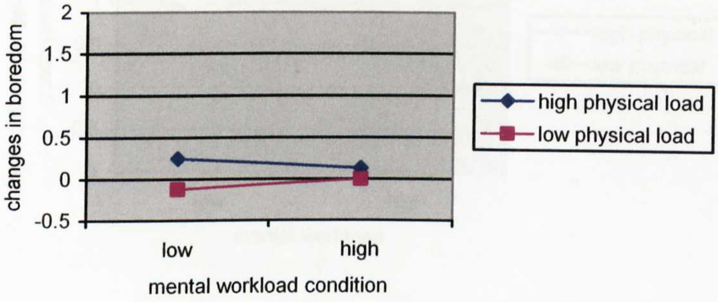


Figure 9.27: Mean change in negative affect as a function of mental and physical workload

9.4.5 Fatigue after-effects

9.4.5.1 Delay until first guess

Mean delay times until the first guess were calculated for pre- and post-CAMS Fault Finding sessions, for four workload conditions. Difference scores were calculated and log transformed scores were then subjected to a 2 x 2 mixed ANOVA. As illustrated in Figure 9.28, there is a trend towards a main effect of mental workload, with times following CAMS being quicker following high mental load. This trend is consistent with the findings presented in Chapter 7, although here, this did not reach significance, $F(1, 11) = 1.49$, $p > 0.05$, $r = .12$. Similarly, there was a slight trend towards a main effect of physical load, with delay times being shorter following high physical load. This was consistent with the findings in Chapter 8, although again was not significant $F(1, 11) < 1$, $p > 0.05$, $r = .03$. There was no significant interaction, $F(1, 11) < 1$, $p > 0.05$, $r = .00$.

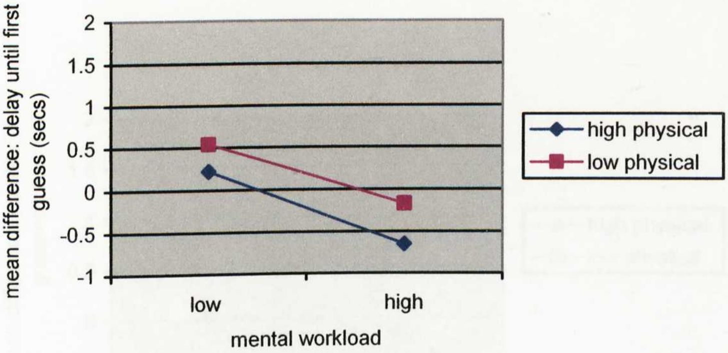


Figure 9.28: Mean difference in delay till first guess as a function of mental and physical workload (a negative score indicates shorter mean delays following the loading task than before)

9.4.5.2 Mean number of guesses

Mean number of guesses were again calculated for the Fault Finding sessions, pre- and post-CAMS, for four levels of workload. Difference scores were calculated and log transformed. Figure 9.29 reveals very little difference between pre- and post-task scores, with slightly fewer guesses following the loading task when physical load was low. As this Figure would suggest, a 2 x 2 mixed ANOVA failed to find a main effect of either mental load, $F(1, 11) < 1$, $p > 0.05$, $r = .01$, or physical load, $F(1, 11) < 1$, $p > 0.05$, $r = .07$, and there was no interaction, $F(1, 11) <1$, $p>0.05$, $r = .11$.

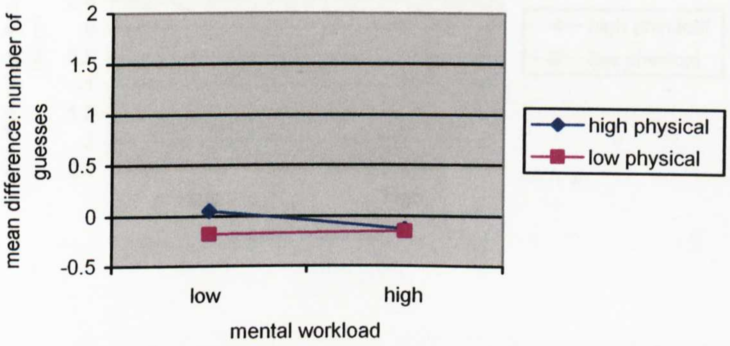


Figure 9.29: Mean number of guesses as a function of mental and physical workload (a negative score indicates fewer mean guesses following the loading task than before)

9.4.5.3 Mean time till solution

Difference scores for mean time till solution were calculated for the four workload conditions and the log transformed scores were subjected to a 2 x 2 mixed ANOVA. Figure 9.30 illustrates a trend towards a main effect of mental workload, with solution times being slower for those following high mental load. However, this effect was again not significant, $F(1, 11) = 1.27$, $p > 0.05$, $r = .00$. Figure 9.30 also illustrates that there was no effect of physical load, $F(1, 11) < 1$, $p > 0.05$, $r = .00$, with there being no difference between the solution times of those following high and low physical load. Figure 9.30 does however illustrate an interesting trend towards an interaction of mental and physical load: While there was no effect of mental workload when physical workload was high, when physical workload was low, low mental workload resulted in faster times to solution, and high mental workload resulted in slower times to solution. Once again, this interesting effect failed to reach significant at the 5% level, $F(1, 11) = 2.75$, $p = 0.08$, $r = .20$, but this finding will be returned to in the discussion.

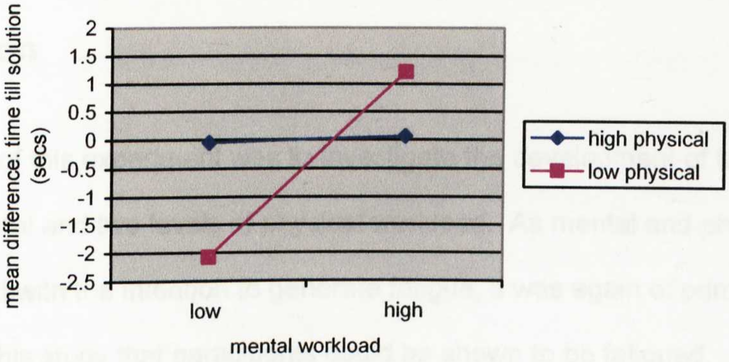


Figure 9.30: Mean difference in time till solution (post minus pre) as a function of workload and effort (a negative score indicates be quicker solution following the loading task than before)

9.5 Discussion

The primary aim of this experiment was to investigate the development of fatigue following two levels of mental and two levels of physical workload. As mental and physical workload were manipulated with the intention to generate fatigue, it was again of primary importance to the validity of this study that participants could be shown to be fatigued.

With regard to mental fatigue, participants were required to complete a complex cognitive task for one hour. This task had been successfully utilised in previous experiments reported here and elsewhere (Sauer, 1997; Hockey, Wastell & Sauer; 1998). This evidence strongly suggests that this task (when incorporating a high number of programmed faults) generates a state of fatigue, in the sense that it produces feelings of tiredness and task aversion. As with previous studies, participants' performance on the primary task was typically very strong, as only a low percentage of time was spent beyond the target state. Consistent with this picture of task engagement was the effect of mental workload on the secondary task of system checks, which was utilised as an indicator of spare capacity. These findings suggest that, under high workload, participants were engaged with the task *and* were working 'hard'. Furthermore, the subjective evaluations of the mental loading task, both concurrently and post-task, support the view that the *high* mental workload manipulation represented a high level of workload and a significantly higher level than the *low* mental workload manipulation.

With regard to physical workload, 70% of VO_2 max was the expert-recommended level for high workload. Furthermore, when utilised in Chapter 8, this level of physical work generated an increase in questionnaire-assessed physical fatigue, a finding which was strongly supported by anecdotal evidence. To confirm the effect of this manipulation in the current experiment, physiological and subjective data were considered. With regard to heart rate, there was a significant effect of physical load during the cycling, with high levels of average heart rate being recorded. Consistent with this was the subjective evaluation of the physical load, both concurrently and post-cycling. Interestingly, manipulation of physical load had a bigger impact on concurrent ratings of overall effort and overall tiredness than the two levels of mental load, which could suggest that broad measures of subjective state are more influenced by physical work than mental work, although this could of course be explained by the alternative proposition that the two levels of physical load represented a greater differential than the two levels of mental load.

In general, this pattern of subjective and objective findings supports the view that the two types of workload were sufficient to be considered as fatiguing. Therefore, having previously identified their separate effects on subjective state and after-effects, what remains is an investigation of their joint effects.

9.5.1 Subjective fatigue

In interpreting the changes in the various aspects of subjective state fatigue, it is important to remember that the main effects of mental and physical load, in the current experiment, each incorporate the combined effects of both levels of the other workload manipulation. Therefore, when considering the main effects of mental load on subjective mental fatigue, changes must be attributed to exposure to both types of work (their separate effects were considered in earlier chapters).

With regard to the effects of mental work, subjective mental fatigue did not change under low mental workload, but significantly increased under high mental workload. This suggests that although carrying out a low level of mental workload whilst cycling was not mentally fatiguing, a high level of mental work and cycling was. Of further interest is the finding that carrying out a high level of mental work did not result in an overall increase in state physical fatigue, but unexpectedly mental workload had a bigger impact at low levels of physical load.

A similar pattern of results was found with regard to the main effects of physical work: Low levels of cycling whilst carrying out a mental task resulted in only very slight increases in physical fatigue, while high level cycling (alongside the mental task) resulted in much larger increases in physical fatigue. But again, the effects of physical work were specific to physical fatigue, as the increases in subjective mental fatigue were almost the same under low and high levels of physical work. This provides further support for the hypothesis that there are distinct mechanisms which generate state mental and state physical fatigue and that participants are differentiating between these resultant states.

Also of interest is the nature of any interaction between mental and physical work on subjective state. While there was no hint of an interaction of mental and physical work on changes in mental fatigue, there was a trend towards an interaction with regard to changes in physical fatigue. While this trend failed to reach significance, the direction of this interaction is particularly interesting, as one might expect the joint effects of mental and physical work to be additive or even multiplicative - i.e. the impact of high mental load would be greater under conditions of high physical load. However, the current findings were to the contrary, as the greatest increase of physical fatigue was within the condition of high physical / low mental load. A possible explanation for this is that participants focused more on the physical discomfort of high physical workload when their mental load was lower.

9.5.2 Fatigue after-effects

As previously discussed, a low-effort Fault Finding strategy is consistent with shorter delays until the first guess, a higher number of average guesses and longer time until solution. With regard to delay until first guess, there were trends towards main effects of both mental and physical load, with delay times being shorter following high workload. While the findings represented only small (and not significant) differences, the trends were consistent with the findings in Chapters 7 and 8. But of particular interest here is the presence or absence of any interaction: However, there was again no hint of interaction of mental and physical fatigue after-effects.

With regard to the average number of guesses, there were surprisingly no main effects and no interaction. Therefore, there was no evidence of the trend towards a low effort strategy present within this variable. More interesting, however, was the mean time till solution. This also surprisingly revealed no main effects of mental and physical workload, which suggests that the quicker average speed of the first guess did not result in costs in time till solution. This is a pattern of results more consistent with the after-effects of only physical load, which again counters the argument that the sources of fatigue are additive or even multiplicative. Of further interest was the strong trend towards an interaction, which failed to reach significance at the 5% level, but yielded a respectable effect size ($r = .20$). The nature of this interaction was that, under high physical workload, there was no mean change in time till solution at either low or high levels of mental workload. When physical workload was low, however, low mental workload resulted in faster times to solution, and high mental workload resulted in slower times to solution. This pattern of results echoes the findings from the changes in subjective physical fatigue, in which the impact of high physical work was greater under low mental work. Both of these findings are consistent with the idea that the development of one type of state fatigue may, counterintuitively, be minimised by distraction of a different type of work: With regard to subjective physical fatigue it is possible that the feelings of aversion to further physical effort were reduced because of the distraction of the high level of mental work. Similarly, the process underlying the adoption of a low effort strategy may have been disturbed by the high level of physical work, which could have operated as a distraction from evaluating subjective state. This is, however, nothing more than a *possible* explanation for these limited trends.

The arguments presented within this, and previous, chapters have sometimes been based on results with fairly low effect sizes, which have sometimes failed to reach significance. There are a number of factors contributing to this: Firstly, these studies would clearly have benefited from a larger sample size. However, because of the necessary participant characteristics and the high level of commitment required to complete a single study, the sample size was inevitably restricted. Secondly, particularly with regard to the after-effects tasks, there was a high degree of within-group variation (or error). A larger sample could have allowed some degree of analysis taking into account individual differences, although this was not possible here.

With these issues in mind, conclusions are put forward only tentatively. However, throughout this thesis the multiple levels of analysis allow a broad view of the complex findings, which do largely offer a coherent picture. It is argued that this strengthens the arguments presented.

Part 4:

General Discussion

Chapter 10

General discussion

10.1 Summary

This chapter provides an overview of the findings from the psychometric and experimental work and a discussion of the conceptual implications of the findings. This is followed by a consideration of the possible theoretical models of fatigue, with particular reference to the advantages and disadvantages of conceptualising fatigue as multidimensional.

10.2 Introduction

This thesis has addressed the conceptual issue of the dimensionality of the fatigue construct. Following the psychometric investigation reported in Part Two, a series of four experiments were undertaken. These were described and discussed in Part Three.

Part Four of the thesis contains only this general discussion chapter and has two general aims. Firstly, to summarise the findings from the psychometric and experimental work and discuss their conceptual implications. Secondly, to consider the possible theoretical models of fatigue, with particular reference to the advantages and disadvantages of conceptualising fatigue as multidimensional.

10.3 Overview of results and conceptual implications

10.3.1 Psychometric work

The preliminary field work carried out in Chapter Three provided support for two distinct strands of enquiry: Some level of differentiation between state and trait fatigue has been implicit within the research literature, particularly with regard to the design of psychometric instruments. As discussed at length in Chapter Four, these have typically varied between a focus on current feelings and a focus on recent (or characteristic) states. However, no evidence was found within this work of an explicit awareness of this differentiation.

This was considered to be an important conceptual distinction, and it was argued that the planned programme of research should be explicit about which aspect (or level) of fatigue it was concerned with. Therefore, the psychometric investigation, which constituted the remainder of Part Two, addressed the issue of dimensionality of fatigue separately for these two levels. This proved to be a valuable approach, as the underlying factor structures of state and trait fatigue were found to be different.

10.3.1.1 Trait fatigue

Trait fatigue was defined as the 'relatively stable individual differences in the predisposition to experience fatigue'. Following on from this definition, it was argued that an investigation of *fatigue predisposition* should be careful not to confound it with the impact of recent activities. This again proved to be a valuable approach, as the validation process revealed that the trait fatigue factors identified were much less closely related than those incorporated in the questionnaire selected to validate these factors against. This was particularly important here as the central concern of the investigation of trait fatigue was the multidimensionality of the construct. Therefore, evidence in support of their independence is significant. Furthermore, it could be argued that this approach has broader implications for the measurement of traits in general. Consider for example the evaluation of trait anxiety: if one has recently undergone a stressful event or period (such as preparing for a PhD viva) one could be expected to rate oneself as 'recently having experienced a high level of anxiety'. This may not be a sensible approach, as it confounds the issues of predisposition with recent events and activities.

With regard to the conceptual implications of the trait fatigue findings, there was strong support for the inclusion of a number of trait-level factors within a model of fatigue. Factor analysis revealed clear and robust distinctions between physical, mental and general fatigue and (at least) two types of sleep-related factors. This suggests there are a number of distinct mechanisms which underlie the development of fatigue. Attempting a detailed description of the nature of these mechanisms is clearly beyond the scope of this thesis. However, there are a number of points worthy of mention. One issue relates to the possible differential stability of these different traits. As an illustration of this consider the distinction between physical fatigue and the sleep-related factors of morning and evening tiredness: The work presented in this thesis supports the view that these can all be considered as types of trait fatigue, as they represent meaningful individual differences in the predisposition to develop tiredness, e.g. some individuals are easily tired by physical work, while others are much less so. However there is an important distinction between these traits: With regard to trait physical fatigue, it could be argued that this factor is partially innate, but is likely to be heavily influenced by current level of individual fitness. Therefore, the threshold for generating a state of fatigue through physical work is (to a substantial degree) flexible.

Conversely, morning and evening tiredness are less likely to be flexible. It is hypothesised that morning and evening tiredness may be related to a biologically pre-set phase length (Horne & Ostberg, 1976): While larks are hypothesized to have a phase length shorter than the 24 hour day, owls are hypothesized to have a phase length longer than 24 hours. This offers a possible explanation of the mechanisms which underlie the individual differences in trait morning and evening tiredness and, if one accepts this proposition, it would follow that these traits would be more predetermined and stable.

As for the possible mechanisms which underlie trait mental fatigue, this is less clear. While planning and the coordination and organization of behaviour have been put forward as processes which breakdown when *in* a state of fatigue (Bartlett, 1943; Broadbent, 1979; & Holding, 1983), this is essentially a different level of question. Here, the issue is the nature of the individual differences which predispose people to a particular threshold of state fatigue. Interestingly, some indirect evidence bearing on this issue comes from the work of Jongman *et al* (in press) who found that low fluid intelligence predicted individuals' shifts to low-effort strategies when fatigued. It is argued that this is because fatigue reduces working memory capacity, which is argued to be associated with lower intelligence. However, while this could explain some level of individual variance, it is likely to be only one of a number of mechanisms involved. Nonetheless, it offers an interesting hypothesis.

Therefore, while there is strong evidence in support of a multidimensional trait fatigue construct, the mechanisms which underlie these individual differences are unclear. It is likely that they are the products of complexly related classes of factors including innate biological differences, experience of and exposure to different fatigue generators and also the cognitive evaluation of the sensory experience. Nonetheless, the multidimensional construct of trait fatigue is an important aspect for consideration in the development of a comprehensive fatigue model.

10.3.1.2 State fatigue

State fatigue was defined as a transitory reaction or process taking place at any given time and level of intensity. While the trait fatigue work described above is interesting, it is state fatigue which has been the focus of so much research and theoretical debate over the past hundred years. The investigation of state fatigue presented in this thesis has provided some insight into the incredible complexity of the theoretical issues involved. While the subjective and psychometric work reported in Part Two did provide some interesting findings to move this discussion forward, the results could be described as 'cloudy'.

In essence, the preliminary interviews provided some limited support for the subjective differentiation between the 'types' of fatigue which follow from different conditions. However, while interviewees were happy to say that these states did differ, they were able to make little meaningful distinction. Common to all experiences was the subjective feeling of tiredness, and each different type of fatigue was associated with different 'correlates' – for example, emotionally demanding circumstances led to the development of tiredness with anxiety. Whereas, physical work led to the development of tiredness along with aching muscles (but typically not anxiety). To consider further the relationships between the potentially different types of fatigue, psychometric methods were again employed. The argument for the adoption of this methodology was that a factor analysis of questionnaire responses should identify whether individuals (both generally and following specific conditions) report differing levels of the fatigue types.

The evidence provided by this psychometric work was mixed. While there was clear support for the differentiation of fatigue from the constructs of boredom and negative affect, it was only when the specific fatigue items were entered into a separate analysis that a forced three factor solution provided any degree of coherence for the three factors of mental, physical and sleep-related fatigue. This suggests that participants were responding to some degree of distinction, but it is unclear whether the level of distinction would be best described as different types of fatigue state *or* that there is only one subjective experience of fatigue, with different correlates depending on the source of the fatigue. The validation process offered little further evidence towards this issue. The correlations within the diary-based FQ factors were much higher than for the trait fatigue, suggesting less independence within the dimensions. Of further interest were the individual differences within these correlations: For some participants there were strong relationships between the three scales, whereas others seemed to be differentiating between the different types of state fatigue. One possible explanation for this is that these two general groups possess a different implicit model of what it is to be mentally or physically fatigued, for example. This is essentially an issue of breadth of definition: When considering their individual levels of physical fatigue, some participants may be incorporating the broader range of correlates which occur following physically demanding work, whereas others may be focusing purely on the their subjective level of tiredness and could therefore be argued to be simply fatigued, i.e. experiencing subjective tiredness and aversion to further effort. An interesting issue then is the level of generality (or specificity) within this aversion to further effort. Evidence of this nature would offer further support for arguments around a unidimensional or multidimensional construct of state fatigue. This was an important aspect of the experimental work.

10.3.2 Experimental work

The main aim of the experimental work was to investigate the impact of different types of workload on subjective state and mental after-effects.

10.3.2.1 Subjective state

This work provided some further evidence to support a differentiation between mental, physical and sleep-related state fatigue, on a subjective level. Following mental work, there were general (although small) increases in physical fatigue and increases in sleep-related fatigue irrespective of the level of workload, whereas mental fatigue increased with increasing levels of mental workload. Also, in line with the predictions of the Compensatory Control model, subjective mental fatigue was found to be particularly sensitive to high levels of workload when combined with a high degree of effort.

Conversely, physical workload had almost no impact on mental and sleep-related fatigue, at any level of workload. Although, physical fatigue increased with increasing levels of physical workload. The impact of separate types of workload suggests that individuals are differentiating, on some level, between the different types of subjective experience. What is unclear from the separate effects, however, is the extent to which they are responding to the wider correlates of this state, or even evaluating their state as being physical or mental simply because of the manipulation they have just completed.

Of further relevance to this question is the impact of concurrent physical and mental loads. There was surprisingly no hint of an additive or multiplicative interaction between changes in subjective mental fatigue and changes in subjective physical fatigue following concurrent manipulations of mental and physical work. This was surprising, as one might expect an interaction if there were some degree of shared mechanisms, even at the subjective level. Again it was surprising that there was a slight (and not significant) trend towards an interaction in the opposite direction for subjective physical fatigue. Rather than an additive or multiplicative model of state fatigue, this could possibly be described as a *masking* model, in which the effects of physical load are greater when mental demands are sufficiently low (possibly) to allow participants to focus on any negative sensory experience. Although it should be noted that this only applied to changes in subjective physical fatigue, which does provide some weight to the argument that these states are, in fact, quite distinct. Nonetheless, this is clearly an interesting area for further research.

10.3.2.2 Performance after-effects

The after-effects task carried out both before and after all workload manipulations provides a different level of evidence to inform these theoretical questions. As expected, strategies employed to carry out this mental task work were found to be sensitive to mental workload and effort. Following high levels of mental work, particularly in the high effort condition, participants' strategies were argued to be consistent with low-effort processing or guessing. However, when investigating the after-effects of physical workload, participants were again quicker to make their first guess following high physical load, but in this manipulation were quicker to find the solution. This suggests that *physically* fatigued participants were simply faster. Therefore, any impact of physical work (at this level) did not transfer to the mental after-effects task. This finding is consistent with the earlier work of Barth, Holding and Stamford (1976) who found that physical work had virtually no effect on a perceptual (risk- and effort-sensitive) task. But, using a physical version of the same task, they found that the effects of one type of physical work did generalise to a broad aversion to further physical effort.

The findings from Experiments Two and Three, taken together, provide good evidence to support the proposition of a multidimensional construct, as this would suggest the mechanisms underlying the development of fatigue following mental work and fatigue following physical work are distinct. However, the nature of any interaction again provides a further level of information: With regard to the delay until first guess, the trends were in line with those of the previous experiments, i.e. both high mental work and high physical work had main effects of reducing the delay until first guess. But there was no evidence of any interaction. With regard to the average number of guesses to find the correct solution, there were no effects of either mental or physical work. However, while there was no main effect of either physical or mental load on time to solution, there was an interesting trend towards an interaction with this variable: Only when physical load was low was there an effect of mental workload. The nature of this (non significant) effect was that, at low physical load, participants were quicker to find a solution when mental workload was also low, and slow to find a solution when mental workload was high. This is very difficult to interpret, but also intriguing if supported by further research. A lower effort strategy would be consistent with a shorter delay to first response, a higher number of mean guesses and a longer time to solution. A high effort strategy would be consistent with a longer time to solution, fewer guesses and a shorter response time. These results are not clearly consistent with either of these findings, but are similar to the performance following just physical work. Furthermore, they are also consistent with the subjective physical fatigue findings following concurrent workload, which suggested that a masking model may offer the best explanation of these results.

10.4 Towards a model of fatigue

The findings presented in this thesis and discussed above do not offer unequivocal support for a multidimensional model of fatigue. However, there is fairly strong and convincing evidence that the fatigue state which develops following different types of work and in different conditions can be distinguished on a subjective level and also that these different types of fatigue have different after-effects.

What are the implications of all of these findings for a model of fatigue? Firstly, any comprehensive model of fatigue should address issues of trait fatigue as well as the more traditional issues of state fatigue. Trait fatigue may be best conceptualised as individual differences in the threshold of the development of fatigue. This concept is clearly multidimensional in nature.

With regard to state fatigue, though, the picture is far more complex. This thesis set out to investigate the relationships between the possible different types of fatigue, with the aim of contributing to a meaningful model or theory of fatigue. As presented here, the findings offer most support for a multidimensional model, with separate mechanisms underlying (at least partially distinct) states.

However, there is an alternative perspective from which these results can be viewed: The *masking model* put forward in Chapter Nine and discussed again above suggests that, at this level, the mechanisms which generate performance-related after-effects do not interact. Further to this, it appears that the development of one type of fatigue may, counterintuitively, be minimised by the distraction of a different type of work. This takes us back full circle to the discussion about what *fatigue* really is, as this supports the conceptualisation of fatigue as a motivational state, which is characterised by a subjective state of tiredness and an aversion to further effort – a 'stop emotion'. This is consistent with some excellent recent research by Matthews and Desmond (2002) who investigated two possible mechanisms for performance impairment in driving, when in a state of mentally induced fatigue. These two mechanisms were (1) a loss of attentional resources and (2) the regulation of effort. In summary, they argued that their results supported the loss of task-directed effort as the mechanism underlying performance impairment, as the performance impairment was lower when the task was more complex. This is clearly at odds with a resource-based or reductionist explanation in which fatigue could be viewed as a reduced quantity of a given resource.

Also consistent with this view of fatigue as a motivational state is the observation by Gaillard (2001) that fatigue is not only determined by the amount of work done, but also what still has to be done. This would again be difficult to explain if we view fatigue from a resource perspective or any other than an emotional/motivational perspective, as the amount of work to be done should not have any influence on the degree of current fatigue if fatigue is a state dependent on available levels of a resource.

So then, given these arguments, should fatigue be considered to be a multidimensional state or a unidimensional state? Essentially, the central issue is one of breadth of definition: If we take a broad view of fatigue and view it as the constellation of effects which follow different conditions and experiences, then we would have clear evidence to support a multidimensional construct. However, there are advantages and disadvantages to this model. Firstly, we return to the problem outlined by Muscio (1921a) and Cameron (1973) (see Part One) who argued that this conceptualisation of fatigue has no explanatory value, but reduces us to simply describing the observed effects. Secondly, there is the problem that we could, in theory, go on to define as many different types of fatigue as there are fatigue-generating conditions – providing little theoretical coherence. However, this approach does have the advantage that there are a number of meaningful and consistent patterns of correlates which follow on from mental work, physical work and sleep deprivation. Of particular current interest is the broad state which follows from high levels of mental work or prolonged mental demands: Consider the excellent recent work for the Dutch fatigue research group headed by Meijman (see van der Linden, Frese & Meijman, 2003; Jongman *et al*, in press; and Schellekens, Sijtsma, Vegter & Meijman, 2000). This work has largely followed in the same vein as the Cambridge Cockpit group (e.g. Bartlett, 1943), and focused on the breakdown in the various mechanisms affected by mental work. This and similar work has cited impairments in planning, coordination and coherence of activity and working memory as possible processes underlying performance effects. This has obvious applied relevance and will continue to inform system design. However, following on from the theoretically coherent work of Bartley and Chute (1947), these effects may be better described as impairments, rather than fatigue proper. Within a model of fatigue, this would then place them as processes which are sensitive to fatigue.

This leaves us with the narrower definition of fatigue as a 'stop emotion'. This is, to some degree, consistent with all the findings presented here, and represents a conceptually clear and coherent construct. However, the proposition that fatigue should be considered as a construct characterised by a subjective state of tiredness and an aversion to further effort does not necessarily imply a unidimensional model. One could argue that the actual feelings of tiredness which follow different fatigue generators are not subjectively the same and that there is more than one feeling of tiredness. The evidence for this proposition is mixed, as interviewees reported that these states were different. This was supported by the subjective findings following the experimental manipulations. However, it is unclear to what extent this can be attributed to broader individual models of fatigue, which incorporate wider 'impairments', or the degree to which they are responding to the overt circumstances which generate the state.

However, with regard to the aversion to further effort, there is a little more information. Following the manipulation of mental and physical work, the aversion to further effort did not generalise from physical work to the investment of further mental effort. This, and the findings of Barth *et al* (1976) suggests that this aspect of fatigue should be viewed as multidimensional as it was only found to generalise within broad parameters, e.g. to any further physical tasks following physical work and any further mental tasks following mental work. The extent and nature of this generalisation presents an interesting range of possibilities for future research.

Therefore, while there is an excellent theoretical justification for adopting a narrow definition of fatigue – this does not preclude a multidimensional construct. Furthermore, although the evidence for this is not comprehensive, there are also practical advantages to this approach: Consider the criticisms of one dimensional fatigue scales (discussed in Chapter Four) which have been argued to be insufficient to capture the full complexity of the subjective experience. Furthermore, multidimensional scales have a diagnostic ability that is important to health professionals, for both assessment and intervention purposes, and in industrial settings, for system design. Moreover, these observations are consistent with the subjective and psychometric work presented within the thesis, which found that the participants and questionnaire respondents consistently differentiated between (what were to them) different types of fatigue. Therefore, the adoption of a narrow, but multidimensional, model has both real psychologically meaning and value in applied settings.

With regard to the taxonomy of this narrow fatigue construct, there is preliminary evidence here to support the inclusion and differentiation of mental, physical and sleep-related dimensions. Although there may be some value in the incorporation of further dimensions, there is no strong evidence for this at present.

REFERENCES

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

Ahsberg, E., Gamberale, F., & Gustaffsson, K. (2000). Perceived fatigue after mental work: an experimental evaluation of a fatigue inventory. *Ergonomics, 43*(2), 252-268.

Aisbitt, B. J. (1998). *Toward the development of a psychometric test of fatigue.* Unpublished dissertation, Hull University, Hull.

Anderson, J. R. (1990). *The adaptive character of thought.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Andreassi, J. L. (1965). Effects of induced muscular tension and auditory stimulation of tachistoscope performance. *Perceptual and Motor Skills, 20,* 829-841.

Annesi, J. J. (2002). Relation of rated fatigue and changes in energy after exercise and over 14 weeks in previously sedentarty women exercisers. *Perceptual and Motor Skills, 95,* 719-727.

Anon. (1946). *The role of fatigue in pilot performance.* (No. Report No. 61): US Civil Aeronautics Administration.

Ash, I. E. (1914). Fatigue and its effect on control. *Archives of Psychology, 4,* 1-61.

Astrand, P. O., & Rodahl, K. (1986). *Textbook of work physiology* (3rd ed). New York: McGraw-Hill.

Baddeley, A. D. (1966). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quarterly Journal of Experimental Psychology, 18*(4), 302-309.

Barmark, J. E. (1940). The time of administration and some effects of 2 grs. of alkaloid caffeine. *Journal of Experimental Psychology, 27,* 690-698.

Barroso, J., & Lynn, M. R. (2002). Psychometric properties of HIV related fatigue scale. *Journal of the association of nurses in AIDS care, 13*(1), 66-75.

Barth, J. L., Holding, D. H., & Stamford, B. A. (1976). Risk versus effort in the assessment of motor fatigue. *Journal of Motor Behaviour, 8,* 189-194.

Bartlett, F. C. (1943). Fatigue following highly skilled work. *Proceedings of the Royal Society, 131*, 247-257.

Bartlett, F. C. (1953). Psychological Criteria of Fatigue. In A. T. Welford, & W.F. Floyd (Ed.), *Fatigue* (pp. 196). London: H.K. Lewis & Co. Ltd.

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

Bates, B. T., Osternig, L. R., & James, S. L. (1977). Fatigue effects in running. *Journal of Motor Behaviour, 9*, 203-207.

Belza, B. L., Henke, C. J., Yelin, E. H., Epstein, W. V., & Gilliss, C. L. (1993). Correlates of fatigue in older adults with rheumatoid arthritis. *Nursing research, 42*(2), 93-99.

Benedict, F. G., and Benedict, C. G. (1933). *Mental effort in relation to gaseous exchange, heart rate and mechanics of respiration.* Washington.

Berlyne, D. E. (1960). *Conflict, arousal and curiosity* (Vol. xii). New York, NY, US: McGraw-Hill.

Bertelson, P., & Joffe, R. (1963). Blockings in prolonged serial responding. *Ergonomics, 6*, 109-116.

Bills, A. D., & Stauffacher, J. C. (1937). The influence of voluntary induced tension on rational problem solving. *Journal of Psychology, 4*, 261-271.

Bills, A. G., & Shapin, M. J. (1936). Mental fatigue under automatically controlled rates of work. *Journal of General Psychology, 15*, 335-347.

Bolger, N., DeLongis, A., Kessler, R. C., & Schilling, E. A. (1989). Effects of daily stress on negative mood. *Journal of Personality and Social Psychology, 57*(5), 808-818.

Borg, G. (1978). Subjective aspects of physical and mental load. *Ergonomics, 21*(3), 215-220.

Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica, 81*, 211-241.

Brehmer, B., & Allard, R. (1991). Dynamic decision-making: the effects of task complexity and feedback delay. In J. Rasmussen, B. Brehmer, & J. Leplat (Ed.), Distributed decision-making: cognitive *models for cooperative work.* New York: John Wiley.

Broadbent, D. E. (1953). Neglect of the surroundings in relation to fatigue decrements in output. In A. T. Welford, & W.F. Floyd (Ed.), *Fatigue* (pp. 196). London: H.K. Lewis & Co. Ltd.

Broadbent, D. E. (1958). *Perception and Communication.* London: Pergammon.

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

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

Brook, R. H., Ware, J. E., & Davis-Avery, A. (1979). Overview of Adult Health Status Measures fielded in Rand's Health Insurance Study. *Medical Care, 17*(7 suppl).

Brookhuis, K. A., & de Waard, D. (2001). Assessment of Drivers' Workload: Performance and Subjective and Physiological Indexes. In P. A. Desmond, & P.A. Hancock (Ed.), *Stress, Workload and Fatigue: Human Factors in Transportation.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Brookings, J. B., Wilson, G.F., & Swain, C.R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biological Psychology, 42*, 361-377.

Brown, I. D., Tickner, A. H., & Simmons, D. C. (1970). Effects of prolonged driving on overtaking criteria. *Ergonomics, 13*, 239-242.

Browne, R. C. (1953). Fatigue, Fact or Fiction? In A. T. Welford, & W.F. Floyd (Ed.), *Fatigue* (pp. 196). London: H.K. Lewis & Co, Ltd.

Brunier, G., & Graydon, J. (1996). A comparison of two methods of measuring fatigue in patients on chronic haemodialysis: Visual analogue vs Likert scale. *International Journal of Nursing Studies, 33*(3), 338-348.

Bulat, P. V. (1971). Evaluation of Methods for the Determination of Factors Inducing Fatigue in Man at Work. *Ergonomics, 14*(1), 43-51.

Caldwell, L. S., & Lyddan, J. M. (1971). Serial isometric fatigue functions with variable intertial intervals. *Journal of Motor Behaviour, 3,* 17-30.

Cameron, C. (1969). *Aircrew fatigue: an empirical and theoretical study.* Unpublished dissertation, Monash University.

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

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

Casali, J. G., & Wierwille, W.W. (1984). On the measurement of pilots perceptual workload: A comparison of assessment techniques addressing sensitivity and intrusion issues. *Ergonomics, 27,* 1033-1050.

Cattell, R. B. (1941). *General Psychology.* Cambridge, Mass.: Sci-Art Publishers.

Cattell, R. B. (1947). Confirmation and clarification of primary personality factors. *Psychometrica, 12,* 197-220.

Chalder, T., Berelowitz, G., Pawlikowska, T., Watts, L., Wessley, S., Wright, D., & Wallace, E. P. (1993). Development of a fatigue scale. *Journal of psychosomatic research, 37*(2), 147-153.

Chambers, E. A. (1961). Industrial fatigue. *Occupational Psychology, 35,* 44-57.

Chiles, W. D. (1955). *Experimental studies of prolonged wakefulness.* Dayton, Ohio: Wright Air Development Centre.

Clarke, H. H. (1958). Physical Fitness Benefits: A Summary of Research. *Education for April, 1958.*

Clarkson-Smith, L., & Hartly, A. A. (1990). Relationships between physical exercise and cognitive abilities in older adults. *Psychology and Aging, 4,* 183-189.

Clough, P. J., Earle. K., & Sewell, D. F. (2002). Mental Toughness: the concept and its measurement. In I. Cockerill (Ed.), *Solutions in Sport Psychology:* Thomson.

Cnossen, F. (1994). *Mental Energy from Energetical and Computational Perspectives.* Groningen: Traffic Research Centre: University of Groningen.

Cnossen, F., Meijman, T., & Rothengatter, T. (2004). Adaptive stategy changes as a function of task demands: A study of car drivers. *Ergonomics, 47*(2), 218-236.

Cohen, S. (1980). After-effects of stress on human performance and social behaviour: a review of research and theory. *Psychological Bulletin, 88*, 82-108.

Cohen, S., & Spacapan, S. (1978). The after-effects of stress: an attentional interpretation. *Environmental Psychology Nonverbal Behaviour, 3*, 43-57.

Cooper, G. E., & Harper, R. P. (1969). *The use of pilot rating in the evaluation of aircraft handling qualities.* (No. NASA TN-D-5153). Moffett Field, CA: Ames Research Center, NASA.

Corwin, W. H., Sandry-Garza, D.L., Biferno, M.H., & Boucek, G.P. (1989). *Assessment of crew workload measurement methods, techniques, and procedures. Vol II.* Guidelines for the use of workload assessment techniques in aircraft certification. (No. WRDC-TR-89-7006). OH: Wright Research and Development Centre, Wright-Patterson Air Force Base.

Craft, D. H. (1983). Effect of prior exercise on cognitive performance tasks by hyperactive and normal young boys. *Perceptual and Motor Skills, 56*, 979-982.

Craft, L. L., & Landers, D. M. (1998). The effects of exercise on clinical depression and depression resulting from mental illness: a meta-analysis. *Journal of Sport and Exercise Psychology, 20,* 339-357.

Craig, A. & Cooper, R. E. (1992). Symptoms of acute and chronic fatigue. In *Handbook of Human Performance* (Vol. 3): Academic Press Ltd.

Crawford, A. (1961). Fatigue and driving. *Ergonomics, 4*, 143-154.

Davey, C. P. (1973). Physical exertion and mental performance. *Ergonomics, 16*, 595-599.

Davis, D. R. (1948). *Pilot Error.* London: HMSO.

Davis, J. M. (1995). Central and peripheral factors in fatigue. *Journal of Sports Sciences, 13*, 49-53.

Dawson, D., & Fletcher, A. (2001). A quantitative model of work-related fatigue: background and definiton. *Ergonomics, 44*(2), 144-163.

de Rijk, A. E., Le Blance, P. M., Schaufeli, W. B., & de Jonge, J. (1998). Active coping and need for control as moderators of the job demand-control model: Effects on burnout. *Journal of Occupational and Organizational Psychology, 71*, 1-18.

de Waard, D. (1996). *The Measurement of Drivers' Mental Workload.* Unpublished PhD Thesis, Groningen, The Netherlands.

DeLongis, A., Folkman, S., & Lazarus, R. S. (1988). The impact of daily stress on health and mood: Psychological and social resources as mediators. *Journal of Personality and Social Psychology, 54*(3), 486-495.

Desmond, P. A., & Hancock, P. A. (2001). Active and Passive Fatigue States. In P. A. Desmond, & P.A. Hancock (Ed.), *Stress, Workload and Fatigue: Human Factors in Transportation.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Desmond, P. A., & Matthews, G. (1997). Implications of task induced fatigue effects for in-vehicle countermeasures to driver fatigue. *Accident Analysis and Prevention, 29*, 513-523.

Devienne, M. F., Audiffren, M., Ripoll, H., & Stein, J. F. (2000). Local muscular fatigue and attentional processes in a fencing task. *Perceptual and Motor Skills, 90*, 315-318.

Dickinson, J., Medhurst, C., & Whittingham, N. (1979). Warm-up and fatigue in skill acquisition and performance. *Journal of Motor Behaviour, 11*, 81-86.

Dinges, D. F., & Kribbs, N. B. (1991). Performing while sleepy: Effects of experimentally-induced sleepiness. In T. H. Monk (Ed.), *Sleep, Sleepiness and Performance*: John Wiley & Sons Ltd.

Dorner, D., & Pfeifer, E. (1993). Strategic Thinking and Stress. *Ergonomics, 36*(11), 1345-1360.

Drew, G. C. (1940). *An experimental study of mental fatigue* (No. 227): Flying Personnel Research Committee.

Drew, G. C. (1979). An Experimental Study of Mental Fatigue. Flying Personnel Research Committee Report, 227. In E. L. Dearnaley & P. B. Warr (Ed.), *Air crew Stress in Wartime Operations.* (pp. 137-177). New York and London: Academic Press.

Duncan, J. (1990). Goal weighting and the choice of behaviour in a complex world. *Ergonomics, 33*(10/11), 1265-1279.

Duncan, J. (1993). Selection of input and goal in the control of behaviour. In A.D. Baddeley & L. Weiskrantz. (Ed.), *Attention, Selection , Awareness and Control: A Tribute to Donald Broadbent* (pp. 53-71). Oxford: Oxford University Press.

Duncan, J. (1995). *Attention, intelligence, and the frontal lobes.* Cambridge, England: Cambridge Medical Research Press.

Easterbrook, J. A. (1959). The effect of emotion on cue utilisation and the organisation of behaviour. *Psychological Review, 66*, 183-201.

Edwards, A. L., & Kilpatrick, F. P. (1948). A technique for the construction of attitude scales. *Journal of Applied Psychology, 32*, 374-384.

Elbel, E. R. (1940). A study of response time before and after strenuous exercise. *Research Quarterley, 11*, 86-95.

Etnier, J. L., Salazar, W., Landers, D. M., Petruzzello, S. J., Han, M., & Nowell, P. (1997). The influence of physical fitness and exercise upon cognitive functioning: A meta-analysis. *Journal of Sport and Exercise Psychology, 19*(3), 249-277.

Eysenck, H. J., & Eysenck, S. B. J. (1964). *Manual of the Eysenck Personality Inventory.* Kent: Hodder and Stoughton.

Eysenck, M. W. (1984). *A handbook of cognitive psychology.* London: Lawrence Erlbaum.

Fairclough, S. (1993). Psychophysiological measures of workload and stress. In S. F. A.M. Parkes (Ed.), *Driving future vehicles.* London: Taylor Francis.

Fairclough, S. (1999). *Mental effort and sustained perceptual-motor performance.* Unpublished Ph.D. Thesis, Loughborough University.

Fairclough, S. (2001). Mental effort regulation and the functional impairment of the driver. In P. A. Desmond, & P. A. Hancock (Ed.), *Stress, Workload and Fatigue: Human Factors in Transportation.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Fairclough, S. H., & Ward, N.J. (1996). A protocol for the assesment of subjective sleepiness. In E. C. Vaya, & J. A. Rothengatter (Ed.), T*raffic and Transport Psychology: Theory and Application.* Amsterdam: Pergamon.

Farh, J. I., & Scott, W. E. (1983). The experimental effects of "autonomy" on performance and self-reports of satisfaction. *Organizational Behavior and Human Performance, 31*, 203-222.

Feather, N. T. (1961). The relationship of persistence at a task to expectation of success and achievement related motives. *Journal of Abnormal and Social Psychology, 63*(3), 552-561.

Feather, N. T. (1963). Persistence at a difficult task with alternative task of intermediate difficulty. *Journal of Abnormal and Social Psychology, 66*(6), 604-609.

Fery, Y. A., Ferry, A., Vom Hofe, A., & Rieu, M. (1997). Effect of physical exhaustion on cognitive functioning. *Perceptual and Motor Skills, 84*, 291-298.

Field, A. (2000). *Discovering statistics using SPSS for Windows advanced techniques for the beginner.* London: Sage.

Figley, C. R. (1995). *Compassion fatigue: Coping with the secondary traumatic stress disorder in those who treat the traumatised.* Philadelphia, PA: Brunner/Mazel, Inc.

Fisk, J. D., Ritvo, P. G., Ross, L., Haase, D. A., Marrie, T. J., & Schlech, W. F. (1994). Measuring the functional impact of fatigue: Initial validation of the fatigue impact scale. *Clinical infectious diseases, 18 (suppl 1)*, S79-83.

Fitts, R. H., & Metzger, J. M. (1993). Mechanisms of muscular fatigue. In J. F. Poortmans (Ed.), *Principles of exercise biochemistry* (2nd ed., Vol. 38, pp. 248-268). Basel: Karger.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin, 51*, 327-358.

Fletcher, A. & Dawson, D. (2001). A quantitative model of work-related fatigue: empirical evaluations. *Ergonomics, 44*(5), 475-488.

Fleury, M., & Bard, C. (1987). Effects of Different Types of Physical Activity on the Performance of Perceptual Tasks in Peripheral and Central Vision and Coincident Timing. *Ergonomics, 30*, 954-958.

Floyd, W. F., & Welford, A. T. (1953). *Symposium on fatigue*. London: H.K.Lewis & Co. Ltd.

Flugel, J. C. (1928). Practice, Fatigue and Oscillation. *British Journal of Psychology, 4*.

Flynn, R. (1972). Numerical performance as a function of prior exercise and aerobic capacity for elementary school boys. *Research Quarterley, 43*, 16-22.

Foltz, E. E., Jung, F. T. & Cisler, L. E. (1944). The effect of some internal factors on human work output and recovery. *American Journal of Physiology, 141*, 641-646.

Frankenhaeuser, M. (1978). *Psychoneuroendocrine approaches to the study of emotion as related to stress and coping*. Paper presented at the Nebraska Symposium on Motivation, US.

Frankenhaeuser, M. (1980). Psychoneuroendocrine approaches to the study of stressful person-environment transactions. In H. Selye. (Ed.), *Selye's Guides to Stress Research* (Vol. 1). New York: Van Nostrand Reinhold.

Frankenhaeuser, M. (1986). A psychobiological framework for research on human stress and coping. In M. H. Appley, & R. Trumbell (Ed.), *Dynamics of stress: Physiological, psychological and social perspectives*. New York: Plenham Press.

Fraser, D. C. (1955). Recent Experimental Work in the Study of Fatigue. *Reprinted from Occupational Psychology, October 1955.*

Frazier, G. (1952). *The development of a scale of subjective feelings of fatigue*. Unpublished M.S.c Thesis, Lafayette, Ind.

Freeman, E. (1939). *Principles of General Psychology*. New York: Henry Halt and Company, Inc.

Froberg, J. E., Karlsson, C. G., Levi, L., & Lidberg, L. (1975). Circadian rhythms of catecholamine excretion, shooting range performance and self-ratings of fatigue during sleep deprivation. *Biological Psychology, 2*, 175-188.

Fry, A. M., & Martin, M. (1996). Fatigue in the chronic fatigue syndrome: A cognitive phenomenon? *Journal of Psychosomatic Research, 41*(5), 415-426.

Funk, S. G., Tornquist, E. M., Champagne, M. T., Copp, L. A., & Wiese, R. A. (1989). *Key aspects of comfort: Management of pain, fatigue and nausea.* New York: Springer Publishing Company.

Gaillard, A. W. K. (2001). Stress, Workload, and Fatigue and Three Biobehavioural States: A General Overview. In P. A. Desmond, & P.A. Hancock (Ed.), *Stress, Workload and Fatigue: Human Factors in Transportation.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Gawron, V. J., French, J., & Funke, D. (2001). An Overview of Fatigue. In P. A. D. P.A. Hancock (Ed.), *Stress, Workload and Fatigue: Human Factors in Transportation.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Gervais, R. (2002). *Patterns of work strain and well being in nurses.* Unpublished PhD Thesis. University of Hull.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory.* Chicago: Aldine.

Glass, D. C., & Singer, J. E. (1972). *Urban Stress.* New York: Academic Press.

Gliner, J. A., Matsen-Twisdale, J. A., Horvath, S. M., & Maron, M. B. (1979). Visual evoked potential and signal detection following a marathon race. *Medicine and Science in Sports, 11*, 155-159.

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 L. K. K.R. Boff, L. Kaufman, J. P. & Thomas (Ed.), *Handbook of perception of human performance* (Vol. II, pp. 41.42-41.49). New York: John Wiley & Sons.

Gopher, D., & Sanders, A. F. (1984). S-Oh-R: Oh stages! Oh resources! In W. Printz, & A. F. Sanders (Ed.), *Cognition and motor processes.* Berlin: Springer-Verlag.

Griffith, J. W., Kerr, W. A., Mayo, T. B. & Topal, J. R. (1950). Changes in subjective fatigue and readiness to work during the eight hour shift. *Journal of Applied Psychology, 34*, 163-166.

Gupta, V. P., Sharma, T. R., & Jaspal, S. S. (1974). Physical activity and the efficiency of mental work. *Perceptual and Motor Skills, 38*, 205-206.

Gutin, B., & DiGennaro, J. (1966). Effect of one-minute and five-minute step-ups on performance of simple addition. *The Research Quarterly, 39*(1).

Gutin, B., & DiGennaro, J. (1968). Effects of a treadmill run to exhaustion on performance of simple addition. *Research Quarterley, 39,* 958-964.

Hackman, J. R., & Oldham, G. R. (1976). Motivation through the design of work: Test the theory. *Organisational Behaviour and Human Performance, 16*, 250-279.

Hamilton, P., Hockey, G. R. J., & Reijman, M. (1977). The place of the concept to activation in human information processing theory: An integrative approach. In S.Dornic (Ed.), *Attention and Performance VI*. New York: Academic Press.

Hammerton, M. (1971). Violent exercise and a cognitive task. *Ergonomics, 14*, 265-267.

Hammerton, M., & Tickner, A.H. (1968). Physical fitness and skilled work after exercise. *Ergonomics, 11*, 41-45.

Hancock, P. A., & Desmond, P. A. (2001). *Stress, Workload and Fatigue: Human Factors in Transportation*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Hardy, G. E., Shapiro, D. A., & Borrill, C. S. (1997). Fatigue in the workforce of national health service trusts: Levels of symptomatology and links with minor psychiatric disorder, demographic, occupational and work role factors. *Journal of Psychosomatic Research, 43*(1), 83-92.

Hart, S. G., & Staveland, L. (1988). Development of the NASA task load index (TLX): Results of empirical and theoretical research. In N. Meshkati, & P.A. Hancock (Ed.), *Human mental workload*. Amsterdam: North-Holland.

Heaton, A. W., & Kruglanski, A. W. (1991). Person perception by introverts and extraverts under time pressure: Effects of need for closure. *Personality and Social Psychology Bulletin, 17*(2), 161-165.

Heckler, C. (1992). Effects of time of posttest after two durations of exercise on speed and accuracy of addition and subtraction by fit and less-fit women. *Perceptual and Motor Skills, 75*, 1059-1065.

Heim, A. W. (1973). *Psychological Testing.* London: Oxford University Press.

Heise, E., Gerjets, P., & Westermann, R. (1997). The influence of a waiting intention on action performance: Efficiency impairment and volitional protection in tasks of varying difficulty. *Acta Psychologica, 97*, 167-182.

Hendy, K. C., Hamilton, K. M., & Landry, L. N. (1993). Measuring subjective workload: When is one scale better than many? *Human Factors, 35*(4), 579-601.

Hendy, K. C., Liao, J., and Milgram, P. (1997). Combining Time and Intensity Effects in Assessing Operator Information-Processing Load. *Human Factors*, 30-47.

Henneman, R. L., & Rouse, W. B. (1984). Measures of human problem solving performance in fault diagnosis tasks. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-14*, 99-112.

Hill, S. G., Iavecchia, H. P., Byer, J. B., Bittner, A. C., Zaklad, A. L., & Christ, R. E. (1992). Comparison of four subjective workload rating scales.. *Human Factors, 34*(4), 429-439.

Hockey, G. R. J. (1984). Varieties of Attentional State: The Effects of Environment. In D. R. Davies, & Parasuraman, R. (Ed.), *Varieties of Attention*: Academic Press.

Hockey, G. R. J. (1986). Changes in operator efficiency as a function of environmental stress, fatigue and circadian rhythms. In K.R. Boff, L. Kaufman & J. Thomas (Ed.), *Handbook of perception and human performance* (Vol. 2, pp. 285-298). Chichester, England: Wiley.

Hockey, G. R. J. (1993). Cognitive-energetical control mechanisms in the management of work demands and psychological health. In A. D. Baddeley & L. Weiskrantz (Ed.), *Attention, selection, awareness and control: A tribute to Donald Broadbent.* Oxford: Oxford University Press.

Hockey, G. R. J. (1996). Energetical-control processes in the regulation of human performance. In. W. Battman & S. Dutke (Ed.), *Process of the molar regulation of behaviour* (pp. 271-287). Scottsdale, A2, USA: Pabst Science Publishers.

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. (1999). Work environments and performance. In N. Chmiel (Ed.), *An introduction to work and organisational psychology.* Oxford: Blackwell.

Hockey, G. R. J. (2000). Environmental Stress: Effects on Human Performance. In *Encylopedia of Stress* (Vol. 2): Academic Press.

Hockey, G. R. J., & Hamilton, P. (1983). The cognitive patterning of stress states. In G. R. J. Hockey (Ed.), *Stress and fatigue in human performance.* Chichester: Wiley.

Hockey, G. R. J., & Meijman, T. (1998). *The construct of Psychological Fatigue: A Theoretical and Methodological Analysis.* Paper presented at the 3rd International Conference on Fatigue and Transportation, Australia.

Hockey, G. R. J., & Wiethoff. M. (1990). Assessing patterns of adjustment to the demands of work. In S. Puglisi-Allegra, & A. Oliverio (Ed.), Psy*chobiology of Stress.* Dordrecht: Kluwer.

Hockey, G. R. J., Gaillard, A. W. K., & Coles, M. G. H. (1986). *Energetics and human information processing.* Dordrecht: Nijhoff Publishers.

Hockey, G. R. J., Maclean, A., & Hamilton, P. (1981). State changes and the temporal patterning of component resources. In J. Long, & A. D. Baddeley. Long (Ed.), *Attention and Performance IX* (pp. 607-621). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Hockey, G. R. J., Payne, R. L., & Rick, J. T. (1996). Intra-individual patterns of hormonal and affective adaptation to work demands: An n = 2 study of junior doctors. *Biological Psychology, 42*(3), 393-411.

Hockey, G. R. J., Wastell, D. G., & Sauer, J. (1998). Effects of Sleep Deprivation and User Interface on Complex Performance: A multilevel Analysis of Compensatory Control. *Human Factors, 40*(2), 233-253.

Hogervorst, E., Reidel, W., Jeukendrup, A., & Jolles, J. (1996). Cognitive performance after strenuous physical exercise. *Perceptual and motor skills, 83*, 479-488.

Holding, D. H. (1983). Fatigue. In G. R. J. Hockey (Ed.), *Stress and Fatigue in Human Performance*. New York: John Wiley & Sons Ltd.

Horne, J. A., & Ostberg, O. (1976). A self assessment questionnaire to determine morningness and eveningness in human circadian rhythms. *International Journal of Chronobiology, 4*, 97-110.

Howell, D. C. (1992). *Statistical methods for psychology*. Boston: PWS-Kent Publishing.

Hueting, J. E., & Sarphati, H. R. (1966). Measuring Fatigue. *Journal of Applied Psychology, 50*(6), 535-538.

Hughes, E. R., Hassoun, J. A., Ward, G. F., & Reub, J. D. (1990). *An assessment of selected workload and situational awareness metrics in a part-mission simulation.* (No. Tech Report ASD-TD-90-5009): Wright-Patterson Air Force Base, O.H: DCS for Integrated Engineering and Technical Management, Aeronautical Systems Division, Air Force System Command.

Humphrey, D. G., and Kramer, A. F. (1994). Toward a Psychophysiological Assessment of Dynamic Changes in Mental Workload. *Human Factors, 36*(1), 3-26.

Humpreys, M. S., & Revelle, W. (1984). Personality, motivation, and performance: a theory of the relationship between individual differences and information processing. *Psychological Review, 91*, 153-184.

Jackson, P. R., Wall, T. D., Martin, R., & Davids, K. (1993). New measures of job control, cognitive demand and production responsibilities. *Journal of Applied Psychology, 78*(5), 753-762.

Jager Adams, M., Tenney, Y. J., & Pew, R. W. (1995). Situation Awarenes and the Cognitive Management of Complex Systems. *Human Factors, 37*(1), 85-104.

Jarrard, L. E. (1960). The role of visual cues in the performance of ergographic work. *Journal of Experimental Psychology, 60*, 57-63.

Jongman, L., Meijman, T., & de Jong, R. (In press). Thinking or doing? Which strategy do people choose when they are mentally fatigued?

Kahneman, D. E. (1973). *Attention and Effort*: Englewood Cliffs. N.J. Prentice Hall.

Kalsbeek, J. W. H., & Ettema, J.H. (1963). Continuous recording of heart rate and the measurement of perceptual load. *Ergonomics, 6*, 306-307.

Kantowitz, B. H. (1978). On the accuracy of speed-accuracy tradeoff. *Acta-Psychologica, 42*(1), 79-80.

Karasek, R. A. (1979). Job Demands, Job Decision Latitude, and Mental Strain: Implications for Job Redesign. *Administrative Science Quarterly, 24*, 288-308.

Kashiwagi, S. (1971). Psychological rating of human fatigue. *Ergonomics, 14*(1), 17-21.

Kersholt, J. H. (1996). The effect of information costs on strategy search in dynamic tasks. *Acta Psychologica, 94*, 273-290.

Kirlik, A. (1993). Modeling Strategic Behaviour in Human-Automation Interaction: Why an "Aid" Can (and Should) Go Unused. *Human Factors, 35*(2), 221-242.

Kline, P. (1986). *A Handbook of Test Construction: Introduction to Psychometric Design*: Methuen & Co Ltd.

Kline, P. (1993). *The Handbook of Psychological Testing*. London: Routledge.

Kogi, K., Saito, Y., & Mitsuhashi, T. (1970). Validity of three components of subjective fatigue. *Journal of Science of Labour, 46*(5), 251-270.

Krupp, L. B., LaRocca, N. G., Muir-Nash, J., & Steinberg, A. D. (1989). Application to patients with multiple sclerosis and systematic lupus erthematosus. *Archives Neurology, 46*, 1121-1123.

Kubitz, K. A., & Pothakos, K. (1997). Does aerobic exercise decrease brain activation? *Journal of sport and exercise psychology, 19*, 291-301.

LaChapelle, D. L., & Finlayson, M. A. J. (1998). An evaluation of subjective and objective measures of fatigue in patients with brain injury and healthy controls. *Brain Injury, 12*(8), 649-659.

Latham, G. P., & Locke, E. A. (1991). Self-regulation through goal setting. *Organizational-Behavior-and-Human-Decision-Processes., 50*(2) Special Issue: Theories of cognitive self-regulation, 212-247.

Lee, K. A., Hicks, G., & Nino-Murcia, G. (1991). Validity and reliability of a scale to assess fatigue. *Psychiatry Research, 36*, 291-298.

Lichtman, D., & Posner, E. G. (1983). The effects of exercise on mood and cognitive functioning. *Journal of Psychosocial Research, 27*, 43-52.

Lorist, M. M., & Tops, M. (2003). Caffeine, fatigue and cognition. *Brain and Cognition, 53*, 82-94.

Lubin, A., Hord, D. J., Tracy, M. L., & Johnson, L. C. (1976). Effects of exercise, bedrest and napping on performance decrement during 40 hours. *Psychophysiology, 13*(4), 334-339.

Lui, Y. (1996). Quantitative assessment of effects of visual scanning on concurrent task performance. *Ergonomics, 39*(3), 382-399.

Lui, Y., & Wickens, C. D. (1994). Mental workload and cognitive task automaticity: an evaluation of subjective and time estimation metrics. *Ergonomics, 37*(11), 1843-1854.

Lundberg, U. (1980). Catecholamine and cortisol excretion under psychologically different laboratory conditions. In E. Usdin, R. Kventriansky & I.J. Kopin (Ed.), Catecholamines and Stress: *recent advances* (pp. 455-460). New York: Elsevier /North-Holland.

Luximon, A., & Goonetilleke, R. S. (2001). Simplified subjective workload assessment technique. *Ergonomics, 44*(3), 229-243.

Lybrand, W. A., Andrews, T. G., & Ross, S. (1954). Systemic fatigue and perceptual organisation. *American Journal of Psychology, 67*, 704-707.

Lyndall, S., & Bicknell, J. (2001). Trauma and the therapist: The experience of therapists working with the penetrators of sexual abuse. *Australasian Journal of Disaster and Trauma Studies, 5*(1).

McDougall, W. (1905) British Journal of Psychology, 1, 435.

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology, 1*, 6-21.

Magnusson, A. E., Nias, D. K. B., & White, P. D. (1996). Is perfectionism associated with fatigue? *Journal of Psychosomatic Research, 41*(4), 377-383.

Malik, M. (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal, 17*, 354-381.

Martinsen, E. W. (1993). Theraputic impications of exercise for clinically anxious and depressed patients. *International Journal of Sport Psychology, 24,* 182-199.

Matthews, G. (1996). *Individual differences in driver stress and performance.* Paper presented at the Human Factors and Ergonomics Society, 40th annual meeting, Santa Monica, CA.

Matthews, G., & Desmond, P. A. (2002). Task-induced fatigue states and simulated driving performance. *The Quarterly Journal of Experimental Psychology, 55A*(2), 659-686.

Matthews, G., & Westerman, S. J. (1994). Energy and attention as predictors of controlled visual and memory search. *Personality and Individual Differences, 17*, 617-626.

Matthews, G., Davies, D. R., Westerman, S. J., & Stammers, R. B. (2000). *Human Performance: Cognition, stress and individual differences*: Psychology Press.

McGlynn, G. H., Laughlin, N. T., & Bender, V. L. (1977). Effect of strenuous to exhaustive exercise on a discrimination task. *Perceptual and Motor Skills, 44*, 1139-1147.

McGlynn, G. H., Laughlin, N. T., & Rowe, V. (1979). The effect of increasing levels of exercise on mental performance. *Ergonomics, 22*(4), 407-414.

McMorris, R., & Graydon, J. (1997). The effect of exercise on cognitive performance in soccer-specific tests. *Journal of sport sciences, 15*, 459-468.

McMorris, T., & Graydon, J. (1996). Effect of exercise on soccer decision-making tasks of differing complexities. *Journal of Human Movement Studies, 30*, 177-193.

McNair, D. M., Lorr, M., & Droppleman, L. F. (1971). *EITS manual for the profile of mood states*. San Diego, CA: Educational and Industrial Testing Service.

McNelly, G. W. (1954). *The development and laboratory validation of a subjective fatigue scale*. Unpublished Dissertation Abstracts 14:2135.

Meek, P. M., Nail, L. M., Barsevick, A., Schwartz, A. L., Stephen, S., Whitmer, K., Beck, S. L., Jones, L. S., & Walker, B. L. (2000). Psychometric testing of fatigue instruments for use with cancer patients. *Nursing Research, 49*(4), 181-190.

Meijman, T. F., Mulder, G., van Dormolen, M., & Cremer, R. (1992). Workload of driving examiners: A psychophysiological field study. In H. Kragt (Ed.), *Enhancing Industrial Performance*. London: Taylor and Francis.

Meyers, C., Zimmerli, W., Farr, S., & Baschnagel, N. (1969). Effects of strenous physical activity on reaction time. *Research Quarterley, 40*(333-337).

Michon, J. A. (1966). Tapping regularity as a measure of perceptual motor load. *Ergonomics, Nine*, 401-412.

Milosevic, S. (1997). Drivers' Fatigue Studies. *Ergonomics, 40*(3), 381-389.

Molloy, R., & Parasuraman, R. (1996). Monitoring an automated system for a single failure: Vigilance and task complexity effects. *Human Factors, 38*(2), 311-322.

Montgomery, G. K. (1983). Uncommon tiredness among college undergraduates. *Journal of consulting and clinical psychology, 51*(4), 517-525.

Moray, N. (1979). Models and measurement of mental workload. In N. Moray (Ed.), *Mental workload: Theory and measurement.* New York: Plenum.

Moray, N. (1982). Subjective mental workload. *Human Factors, 24*(1), 25-40.

Moray, N., Dessouky, M. I., & Kijowski, B. A. (1991). Strategic Behaviour, Workload, and Performance n Task Scheduling. *Human Factors, 33*(6), 607-629.

Morgan, B. B., & Alluisi, E.A. (1972). Synthetic work: Methodology for assessment of human performance. *Perceptual Motor Skills, 35*, 835-845.

Morgan, W. P. (1997). *Physical activity and mental health* (Vol. xv). Philadelphia, PA, US: Taylor and Francis.

Morrison, D. L., & Duncan, K. D. (1988). Strategies and tactics in fault diagnosis. *Ergonomics, 31*(5), 761-784.

Morriss, R. K., Wearden, A. J., & Battersby, L. (1997). The relation of sleep difficulties to fatigue, mood and disability in chronic fatigue syndrome. *Journal of Psychosomatic Research, 42*(6), 597-605.

Morriss, R. K., Weardon, A.J., & Mullis, R. (1998). Exploring the validity of the Chalder fatigue scale in chronic fatigue syndrome. *Journal of Psychosomatic Research, 45*(5), 411-417.

Mosso, A. (1890). Les lois de la fatigue etudiées dans les muscles de l'homme. *Archives of Italian Biology, 13*, 123-186.

Muckler, F. A., & Seven, S. A. (1992). Selecting performance measures: 'Objective' versus 'subjective' measurement. *Human Factors, 34*, 441-455.

Mulder, G. (1980). *The heart of mental effort.* University of Groningen, Groningen.

Mulder, G. (1986). Mental effort and its measurement. In G.R.J. Hockey, A. W. K. Gaillard, & M. G. H. Coles (Ed.), *Energetics and human information processing.* (pp. 175-198). Dordrecht: Martinus Nijhoff Publishers.

Muraven, M., Tice, D. M., & Baumeister, R. F. (1998). Self-control as a limited resource: Regulatory depletion patterns. *Journal of Personality and Social Psychology, 74*(3), 774-789.

Murray, E. A., and Caldwell, B. S. (1996). Human performance and control of multiple systems. *Human Factors, 38*(2), 323-329.

Muscio, B. (1921a). Is a fatigue test possible? *British Journal of Psychology, 12*, 13-46.

Muscio, B. (1921b). Feeling-tone in industry. *British Journal of Psychology, 12*, 150-162.

Neville, K. J., Bisson, R. U., French, J., & Boll, P. A. (1994). Subjective fatigue of C-141 aircrew during operation desert storm. *Human Factors, 36*(2), 339-349.

Norman, D. A., & Bowbrow, D.G. (1975). On data-limited and resource-limited processes. *Cognitive psychology, 7*, 44-64.

Nygren, T. E. (1991). Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload. *Human Factors, 31*(1), 17-33.

O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology. In K.R. Boff., L. Kaufman., & J.P. Thomas (Ed.), *Handbook of perception and human performance.* (Vol. II, pp. 41/41-42/49). New York: Wiley.

Offner, M. (1911). *Mental Fatigue.* Baltimore: Warwick & York.

Okuyama, T., Akechi, T., Kugaya, A., Okamura, H., Shima, Y., Maruguchi, M., Hosaka, T., & Uchitomi, Y. (2000). Development and validation of the Cancer fatigue scale: A brief three-dimensional self rating scale for assessment of fatigue in cancer patients. *Journal of Pain and Symptom Management, 19*(1), 5-14.

Paas, E. G. W. C., & Adam, J. J. (1991). Human information processing during physical exercise. *Ergonomics, 34,* 1385-1397.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse and abuse. *Human Factors, 39*(2), 230-253.

Parasuraman, R., & Hancock, P. A. (2001). Adaptive control of mental workload. In P. A. Desmond &. P.A. Hancock (Ed.), *Stress, Workload and Fatigue: Human Factors in Transportation.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Parasuraman, R., Mouloua, M., & Molloy, R. (1996). Effects of adaptive task allocaiton on monitoring of automated systems. *Human Factors, 38*(4), 665-679.

Payne, J. W. (1976). Task complexity and contingent processing in decision-making: information search and protocol analysis. *Organisational Behaviour and Human Performance, 16*, 366-387.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker.* New York: Cambridge University Press.

Pearson, P. G., & Byars, G. E. (1956). *The development and validation of a checklist for measuring subjective fatigue* (No. 56-115). Texas: School of Aviation. USAF, Randolf.

Pearson, R. G. (1957). Scale analysis of a fatigue checklist. *Journal of Applied Psychology, 41*(3), 186-191.

Petruzzello, S. J., Landers, D. M., Hatfield, B. D., Kubitz, K. A., & Salazar, W. (1991). A meta-analysis on the anxiety reducing aspects of acute and chronic exercise: outcomes and mechanisms. *Sports Medicine, 11,* 42-182.

Piper, B. F., Lindsey, A. M., Dodd, M. J., Ferketich, S., Paul, S. M., & Weller, S. (1989). The development of an instrument to measure the subjective dimension of fatigue. In S. G. Funk (Ed.), *Key aspects of comfort: Management of pain, fatigue and nausea.* New York: Springer Publishing Company.

Plante, T. T., & Rodin, J. (1990). Physical fitness and enhanced psychological health. *Current Psychology: Research and Reviews, 9*(1), 3-24.

Poffenberger, A. T. (1928). The effects of continuous work upon output and feelings. *Journal of Applied Psychology, 12*(5), 459-467.

Pribram, K. H., & McGuiness, D. (1975). Arousal, activation, and effort in the control of attention. *Psychological Review, 82*(2), 116-149.

Rabbit, P. A. (1981). Sequential reactions. In D. H. Holding (Ed.), *Human Skills.* Chichester: John Wiley & Sons.

Raby, M., & Lee, J. D. (2001). Fatigue and workload in the maritime industry. In P. A. Desmond & P.A. Hancock (Ed.), *Stress, Workload and Fatigue: Human Factors in Transportation.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Reason, J. (1990). *Human Error.* Cambridge: University Press.

Rehman, J. T., Stein, E.S., & Rosenberg, B.L. (1983). Subjective pilot workload assessment. *Human Factors, 25*(3), 297-307.

Reid, G., & Nygren, T. (1988). The subjective workload assessment technique: a scaling procedure for measuring mental workload. In N. Meshkati & P. Hancock (Ed.), *Human Mental Workload* (pp. 185-218). Amsterdam: Elsevier.

Rhoten, D. (1982). Fatigue and the post surgical patient. In C. M. Norris (Ed.), *Concept clarification in nursing*. Rockville, MD: Aspen Systems Corp.

Rivers, W. H. R. (1896). On Mental Fatigue and Recovery. *United Kingdom, 13,* 525-529.

Rosa, R. R. (2001). Examining work schedules for fatigue: It's not just hours of work. In P. A. Desmond & P.A. Hancock (Ed.), *Stress, Workload and Fatigue: Human Factors in Transportation*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Rosenfeld, O., Tenenbaum, G., Rushkin, H., & Halfons, S. (1989). The effect of physical training on objective and subjective measures of productivity and efficiency in industry. *Ergonomics, 32*(8), 1019-1028.

Rotton, J., Olszewski, D., Charleton, M., & Soler, E. (1978). Loud speech, conglomerate noise, and behavioral aftereffects. *Journal of Applied Psychology, 63*(3), 360-365.

Rouse, W. B. (1978). Human problem solving performance in a fault diagnosis task. *IEE transactions on systems, man, and cybernetics, SMC-8,* 258-271.

Rouse, W. B., & Hunt, R.M. (1988). Human problem solving in fault diagnosis tasks. In W. B. Rouse (Ed.), *Advances in Man Machine Systems Research*. New York: JAI Press.

Ryan, A. H., & Warner, M. (1936). The effect of automobile driving on the reactions of the driver. *American Journal of Psychology, 48,* 403-421.

Ryan, T. A. (1944). Varieties of fatigue. *American Journal of Psychology, 57,* 565-569.

Saito, Y., Kogi, D., & Kashiwagi, S. (1970). Factors underlying subjective feelings of fatigue. *Journal of Science of Labour, 46*(4), 205-224.

Sauer, J. (1997). *Human skill maintenance in complex work environments: Applications to extended space flight.* Unpublished Ph.D. Thesis, University of Hull, Hull.

Sauer, J., Wastell, D.G., Hockey, G.R.J., & Earle, F. (2003). Effects of occasional nightwork on complex task performance. *Human Factors*, 45 (4), 657-669.

Schellekens, J. M. H., Sijtsma, G.J., Vegter, E., & Meijman, T. (2000). Immediate and long delayed after-effects of long lasting mental demands. *Biological Psychology, 53*(1), 37-56.

Schonpflug, W. (1983). Coping efficiency and situational demands. In G. R. J. Hockey. (Ed.), *Stress and Fatigue in Human Performance.* Chichester: Wiley.

Schwab, R. S. (1953). Motivation in measurements of fatigue. In A. T. Welford (Ed.), *Fatigue* (pp. 143-148). London: H.K. Lewis & Co. Ltd.

Schwartz, A., & Meek, P. (1999). Additional construct validity of the Schwartz cancer fatigue scale. *Journal of Nursing Measurement, 7*(1), 35-45.

Shallice, T., & Burgess, P. (1993). Supervisory control of action and thought selection. In A. D. Baddeley & L. Weiskrantz (Ed.), *Attention: Selection, awareness, and control: A tribute to Donald Broadbent.* (Vol. XV, pp. 436 pp). New York, NY, US: Clarendon Press/Oxford University Press.

Shiffrin, R. M. & Schneider, W. (1977). Controlled and automatic information processing: II. Perceptual Learning, Automatic attending and a general theory. *Psychological Review* (84), 127-190.

Shingledecker, C. A., & Holding, D.H. (1974). Risk and effort measures of fatigue. *Journal of Motor Behaviour, 6*, 17-25.

Simonson, E., & Weiser, P.C. (1976). *Psychological aspects and physiological correlates of work and fatigue.* Illinois: Charles C Thomas Publisher.

Sjoberg, H. (1980). Physical fitness and mental performance during and after work. *Ergonomics, 23,* 977-995.

Smets, E. M. A., Garssen, B., Bonke, B., & De Haes, J.C.J.M. (1995). The Multidimensional fatigue inventory (MFI) Psychometric qualities of an instrument to assess fatigue. *Journal of Psychosomatic Research, 39*(5), 315-325.

Smith, A. P. (1992). Chronic fatigue syndrome and performance. In *Handbook of Human Performance* (Vol. 2). London: Academic Press.

Spaeth, R. A. (1920). *Industrial education.* New York: The society of industrial engineers.

Sperandio, J. C. (1971). Variation of operator's strategies and regulating effects on workload. *Ergonomics, 14*(5), 571-577.

Sperandio, J. C. (1978). The regulation of working methods as a function of workload among air traffic controllers. *Ergonomics, 211*, 195-202.

Spielberger, C. D. (1983). *Manual for the state trait anxiety inventory.* Palo Alto: Consulting Psychologists Press, Inc.

Stamford, B. A. (1976). Validity and reliability of subjective ratings of perceived exertion during work. *Ergonomics, 19*(1), 53-60.

Stauffacher, J. C. (1937). The effects of induced muscular tension on various phases of the learning process. *Journal of Experimental Psychology, 21*, 26-46.

Sternberg, S. (1969). The discovery of processing stages: extention of Donders' method. *Acta Psychologica, 30*, 276-315.

Stockfelt, T. (1973). *Relations between physiological exertion and mental performance.* (Report from Institute of Education): University of Stockholm.

Stuss, D. T., Shallice, T., Alexander, M. P., & Picton, T. W. (1996). A multidisciplinary approach to anterior attentional functions. In J. Grafman, & K.J. Holyoak (Ed.), *Structure and functions of the prefrontal cortex* (Vol. 769, pp. 191-211). New York: NY Academy of Sciences.

Styles, E. A. (1997). *The Psychology of Attention*: Psychology Press Ltd.

Tabachnick, B. G., & Fidell, L.S. (1996). *Using Multivariate Statistics* (Third ed.). New York: Harper Collins.

Tafalla, R. J., & Evans, G.W. (1997). Noise, physiology and human performance: The potential role of effort. *Journal of Occupational Heal Psychology, 2*(2), 148-155.

Takakuwa, E. (1962). The function of concentration maintenance (TAF) as an evaluation of fatigue. *Ergonomics, 5*, 37-49.

Takakuwa, E. (1977). The functioning of maintaining concentration (TAF): an approach to the evaluation of mental stress. In R. R. Mackie (Ed.), *Vigilance: Theory, Operational Relevance and Physiological Correlates.* (pp. 217-238). Chichester: Wiley.

Tattersall, A. J. (1999). Workload and task allocation. In N. Chmiel (Ed.), *Work and organisational psychology* (pp. 181-274). Oxford: Blackwall.

Tattersall, A. J., & Foord, P.S. (1996). An experimental evaluation of instantaneous self-assessment as a measure of workload. *Ergonomics, 39*(5), 740-748.

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.

Tattersall, A. J., & Rowe, S.J. (1994). Exercise, effort and cognitive performance. In *Contemporary Ergonomics*. London: Taylor Francis.

Teichner, W. H. (1968). Interaction of behavioural and physiological stress reactions. *Psychological Review, 75*, 51-80.

Tepas, D. I., & Price, J.M. (2001). What is stress and what is fatigue? In P. A. Desmond & P.A. Hancock (Ed.), *Stress, Workload and Fatigue: Human Factors in Transportation*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Thorndike, E. (1900). Mental fatigue. *Psychology Review, 7*, 466-482.

Thorndike, E. L. (1917). The curve of work and the curve of satisfyingness. *Journal of Applied Psychology, 1*, 265-267.

Tiesinga, L. J., Dassem, T. W. N., & Halfen, R. J. G. (1998). DUFS and DEFS: Development, reliability and validity of the Dutch Fatigue Scale and the Dutch Exertion Fatigue Scale. *International Journal of Nursing Studies, 35*(1-2), 115-123.

Tiesinga, L. J., Dassen, W. N., Halfens, R. J. G., & van den Heuvel, W. J. A. (1999). Factors related to fatigue; priority of interventions to reduce or eliminate fatigue and the exploration of a multidisciplinary research model for further study of fatigue. *International Journal of Nursing Studies, 36*, 265-280.

Tomporowski, P. D., & Ellis, N. R. (1986). Effects of exercise on cognitive processes: A review. *Psychological Bulletin, 99*(3), 338-346.

Tomporowski, P. D., Ellis, N. R., & Stevens, R. (1987). The immediate effects of strenuous exercise on free-recall memory. *Ergonomics, 30*(1), 121-129.

Tsaneva, N., & Markov, S. (1971). A model of fatigue. *Ergonomics, 14*, 11-16.

van der Linden, D., Frese, M., & Meijman, T. F. (2003). Mental fatigue and the control of cognitive processes: effects on perseveration and planning. *Acta Psychologica, 113*, 46-65.

Veltman, J. A., & Gailliard, A. W. K. (1996). Physiological indices of workload in a simulated flight task. *Biological Psychology, 42*(Special Issue: Psychophysiology of workload 3), 323-342.

Verbrugge, L. M. (1980). Health Diaries. *Medical Care, 18*(1).

Verwey, W. B. (2000). On-line driver workload estimation. Effects of road situation and age on secondary task measures. *Ergonomics, 43*(2), 187-209.

Verwey, W. B., and Veltman, H. A. (1996). Detecting Short Periods of Elevated Workload: A Comparision of Nine Assessment Techniques. *Journal of Experimental Psychology: Applied, 2*(3), 270-285.

Vidulich, M. A., & Tsang, P. S. (1987). *Absolute magnitude estimates and relative-judgement approaches to subjective workload assessment.* Paper presented at the Human Factors Society 31st Annual Meeting, Santa Monica, CA.

Vincente, K. J., Thornton, D. C., & Moray, N. (1987). Spectral analysis of sinus arrhythmia: A measure of mental effort. *Human Factors, 29*(2), 171-182.

Wall, T. D., Jackson, P. R., Mullarkey, S., & Parker, S. K. (1996). The demands-control model of job strain: A more specific test. *Journal of Occupational and Organisational Psychology, 69*, 153-166.

Warren, N., & Clark, B. (1937). Blocking in mental and motor tasks during a 65-hour vigil. *Journal of Experimental Psychology, 21*, 97-105.

Watson, D., Clark, L.A., & Tellegan, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scale. *Personality and Social Psychology, 54*, 1063-1070.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*(2), 219-235.

Webster, D. M., Richter, L., & Kruglanski, A. W. (1996). On leaping to conclusions when feeling tired: Mental fatigue effects on impressional primacy. *Journal of Experimental Social Psychology, 32*, 181-195.

Weingarten, G., & Alexander, J. F. (1970). Effects of physical exertion on mental performance of college males of different physical fitness levels. *Perceptual and Motor Skills, 31*, 371-378.

Weinland, J. D. (1927). The variability of performance in the curve of work. *Archives of Psychology, 87*.

Welford, A. T. (1953). The psychologists problem in measuring fatigue. In A. T. Welford & W. F. Floyd (Ed.), *Fatigue* (pp. 196). London: H. K. Lewis & Co. Ltd.

Welford, A. T., Brown, R. A. & Gabb, J. E. (1950). Two experiments on fatigue as affecting skilled performance in civilian aircrew. *British Journal of Psychology, 40*, 195-211.

Wells, F. L. (1912). The question of association types. *Psychological Review, 19*(4), 253-270.

Wessley, S., & Powell, R. (1989). Fatigue syndromes: a comparison of chronic "postviral" fatigue with neuromuscular and affective disorders. *Journal of Neurology, Neurosurgery and Psychiatry, 52*, 940-948.

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin, 120*(2), 272-292.

Whiting, H. F., & English, H. B. (1925). Fatigue tests and incentives. *Journal of Experimental Psychology, 8*, 33-49.

Wickens, C. D. (1986). Processing resources in attention. In R. Parasuraman & D. R. Davies (Ed.), *Varieties of attention* (pp. 63-102). New York: Wiley.

Wickens, C. D. (1992). *Engineering psychology and human performance.* New York: Harper Collins.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (Third edition ed.). Upper Saddle River, NJ: Prentice Hall.

Wickens, C. D., Kramer, A. F., & Donchin, E. (1984). The event-related potential as an index of the processing demands of a complex target acquisition task. *Annals of the New York Academy of Sciences, 425*, 295-299.

Wierwille, W. W., & Casali, J. G. (1983). *A validated rating scale for global mental workload measurement applications.* Paper presented at the Proceedings of the Human Factors Society 27th annual meeting, Santa Monica: CA.

Wierwille, W. W., & Eggemeier, F. T. (1993). Recommendations for mental workload measurement in a test and evaluation environment. *Human Factors, 35*(2), 263-281.

Wiggers, C. J. (1944). *Physiology in Health and Disease.* (4th ed.). Philadelphia: Lea & Febiger.

Wilkinson, R. T. (1962). Muscle tension during mental work under sleep deprivation. *Journal of Experimental Psychology, 64*, 565-571.

Wilson, G. F. (1992). Applied use of cardiac and respiration measures: Practical considerations and precautions. *Biological psychology, 34*, 163-178.

Wilson, G. F., & Eggemeier, F. T. (1991). Psychophysiological assessment of workload in multi-task environments. In D. L. Damos (Ed.), *Multiple task performance* (pp. 329-360). London: Taylor & Francis.

Wilson, G. F., & Fullenkamp, P. (1991). *A comparison of pilot and WSO workload during training missions using psychophysiological data.* Paper presented at the Proceedings of the Western European Associated for Aviation Psychology: Stress and error in aviation, Brighton, England.

Wilson, G. F., & O'Donnell, R. D. (1988). Measurement of operator workload with the neurophysiological workloads test battery. In N. Meshkati & P.A. Hancock (Ed.), *Human mental workload* (pp. 63 to 100). Amsterdam: North Holland.

Winstead-Fry, P. (1998). Psychometric Assessment of Four Fatigue Scales with a Sample of Rural Cancer Patients. *Journal of Nursing Measurement, 6*(2).

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* (Vol. 84). United Kingdom: Nature Publishing Group.

Yoshitake, H. (1971). Relations between the symptoms and the feeling of fatigue. *Ergonomics, 14*(1), 175-186.

Yoshitake, H. (1978). Three characteristic patterns of subjective fatigue symptoms. *Ergonomics, 21*, 231-233.

Young, P. T. (1936). *Motivation of Behaviour.* New York: John Wiley & Sons, Inc.

Zapf, D., Dormann, C., & Frese, M. (1996). Longitudinal studies in organizational stress research: A review of the literature with reference to methodological issues. *Journal of Occupational Health Psychology, 1*(2), 145-169.

Zeitlin, L. R. (1995). Estimates of Driver Mental Workload: A Long-Term Field Trial of Two Subsidiary Tasks. *Human Factors, 37*(3), 611-621.

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica, 67*, 361-370.

Zijlstra, F. R. H. (1993). *Efficiency in work behaviour. A design approach for modern tools.* Unpublished PhD Thesis, Delft University of Technology, Delft, The Netherlands.

Zinchenko, V. P., Leonova, A. B., & Strelkov, Y. K. (1985). *The psychometrics of fatigue.* London: Taylor & Francis.

Appendices

Appendix 1: Trait Scale (1)

GENERAL TIREDNESS QUESTIONNAIRE 1

name **sex** **age**
occupation

The following list of statements has been compiled from comments people have made about their general feelings of tiredness and energy in different circumstances. Please read each one carefully and indicate how much it applies to your normal everyday state - that is, **how you are generally**. *Circle one of the numbers on each line:* ***1*** *= disagree strongly;* ***2*** *= disagree moderately;* ***3*** *=disagree slightly;* ***4*** *= agree slightly;* ***5*** *= agree moderately;* ***6*** *= agree strongly – i.e. the higher the number the more you agree that this statement applies to the way you generally feel.*

Note: Although some of the statements are similar to others please respond to each one separately. Go through the list from beginning to end without leaving gaps or going back to change responses. Of course, there are no right answers – we are interested only in measuring the variety of experiences that people have in relation to feelings of tiredness and energy across a wide range of situations. A small number of items may not apply to everyone. If an item does not apply to you at all, please write N/A next to it.

Thinking of how I am generally...	**Disagree**					**Agree**
I enjoy hard physical work	1	2	3	4	5	6
I get tired for no reason	1	2	3	4	5	6
I am too tired to think for long periods	1	2	3	4	5	6
I feel energetic over the whole day	1	2	3	4	5	6
I find physical activity very hard	1	2	3	4	5	6
I can resist external distraction and maintain concentration on what I am doing	1	2	3	4	5	6
Listening to other people's problems leaves me feeling drained	1	2	3	4	5	6
I feel sleepy at work in the afternoon	1	2	3	4	5	6
Winter weather make me feel depressed and weary	1	2	3	4	5	6
After a good night's sleep I feel revived	1	2	3	4	5	6
If I am tired a quick nap always perks me up	1	2	3	4	5	6
I have difficulty remaining alert on long journeys	1	2	3	4	5	6
When I have a cold I feel lethargic	1	2	3	4	5	6
Missing meals leaves me feeling tired	1	2	3	4	5	6
I wake up feeling tired	1	2	3	4	5	6
Emotional conflicts with colleagues or friends leave me feeling drained	1	2	3	4	5	6
If I am feeling tired during work I prefer doing less demanding tasks	1	2	3	4	5	6
I get worn out easily by any prolonged mental activity	1	2	3	4	5	6
I expect to feel very lively	1	2	3	4	5	6
Working without a break makes me feel weary	1	2	3	4	5	6
I am not a lively person	1	2	3	4	5	6
I feel worn out after exerting myself physically	1	2	3	4	5	6
I have trouble sleeping at night	1	2	3	4	5	6
I get tired when I need to concentrate hard for long periods	1	2	3	4	5	6
Being physically active makes me mentally alert	1	2	3	4	5	6

Thinking of how I am generally...	**Disagree**					**Agree**
I have no trouble remaining alert during the working day	1	2	3	4	5	6
I get tired when it's really hot	1	2	3	4	5	6
I don't feel like making much of an effort when I am tired	1	2	3	4	5	6
If I stay in for the evening I might feel sleepy	1	2	3	4	5	6
I get up early in the morning	1	2	3	4	5	6
I think of myself as having bags of energy	1	2	3	4	5	6
My capacity for concentration falls off later in the day	1	2	3	4	5	6
I like to take naps during the day	1	2	3	4	5	6
I am unable to sustain a high level of mental effort for long periods	1	2	3	4	5	6
If I am doing something when tired I tend to skip less important details	1	2	3	4	5	6
I take a long time to get going in the morning	1	2	3	4	5	6
I am able to concentrate for long periods without lapses of attention	1	2	3	4	5	6
I only feel alert if I eat at the right times	1	2	3	4	5	6
Having a hot bath makes me feel sleepy	1	2	3	4	5	6
My current sleep pattern works well for me	1	2	3	4	5	6
Strenuous physical work exhausts me	1	2	3	4	5	6
I come across to others as a bit lethargic	1	2	3	4	5	6
I get tired from having to deal with problems or things going wrong	1	2	3	4	5	6
I have great capacity for dealing with emotional problems	1	2	3	4	5	6
I feel tired at the end of a full working day	1	2	3	4	5	6
I can sustain attention even on quite monotonous tasks	1	2	3	4	5	6
I am full of energy first thing in the morning	1	2	3	4	5	6
Physical work leaves me feeling worn out	1	2	3	4	5	6
When I'm hungry I feel drained of energy	1	2	3	4	5	6
After waking I feel sleepy for some time	1	2	3	4	5	6
I get tired when I have been trying to cope with a stressful problem	1	2	3	4	5	6
When I am tired I tend to cut a few corners to get things done	1	2	3	4	5	6
I find other people's problems quite wearing	1	2	3	4	5	6
I have trouble falling asleep at night	1	2	3	4	5	6
I lose interest if what I have to do is not mentally stimulating	1	2	3	4	5	6
My overall level of vitality is low	1	2	3	4	5	6
However tired I am I feel refreshed after sleep	1	2	3	4	5	6
I go to bed early	1	2	3	4	5	6
If there's not much of interest going on in my surroundings I get bored	1	2	3	4	5	6
Prolonged mental activity wears me out	1	2	3	4	5	6
I wake up in the morning before I need to get up	1	2	3	4	5	6
I feel sleepy after I have had a few drinks (of alcohol)	1	2	3	4	5	6
I enjoy taking part in being physically active	1	2	3	4	5	6
Long train journeys leave me feeling sluggish	1	2	3	4	5	6
Sport and exercise make me feel tired	1	2	3	4	5	6
Naps are a good way to keep me feeling fresh	1	2	3	4	5	6
Eating a big meal makes me feel sleepy	1	2	3	4	5	6
I feel tired in the evenings after working all day	1	2	3	4	5	6
Working under pressure makes me feel worn out	1	2	3	4	5	6
Driving for several hours without a break makes me tired	1	2	3	4	5	6

	Disagree					Agree
Thinking of how I am generally...						
Spending time with friends is invigorating	1	2	3	4	5	6
If I have a difficult piece of work to carry out I prefer to do it in the afternoon	1	2	3	4	5	6
If I can I like to have a nap in the middle of the day	1	2	3	4	5	6
I tend to feel a bit sleepy during the afternoon	1	2	3	4	5	6
I recover quickly from hard physical work	1	2	3	4	5	6
I feel very lively in the mornings	1	2	3	4	5	6
Repetitive work makes me feel weary	1	2	3	4	5	6
When I have to get up early I feel tired all day	1	2	3	4	5	6
I feel more lively when I am out of doors	1	2	3	4	5	6
I am happy to slump in front of the TV for the evening	1	2	3	4	5	6
I feel alert most of the time	1	2	3	4	5	6
When I have over-slept I wake up feeling more tired than usual	1	2	3	4	5	6
Doing nothing for a long time makes me feel weary	1	2	3	4	5	6
Left to myself I tend to drop off in the evenings	1	2	3	4	5	6
I feel rested after a long train journey	1	2	3	4	5	6
When I stay up very late to do some work I feel terrible all the next day	1	2	3	4	5	6
If I lose sleep I find myself dropping off during the day	1	2	3	4	5	6
A cup of coffee makes me feel more alert	1	2	3	4	5	6
I feel most energetic around the middle of the day	1	2	3	4	5	6
I like the feeling I get from doing hard physical work	1	2	3	4	5	6
I can stay up late without any ill effects	1	2	3	4	5	6
A few drinks (of alcohol) makes me lively and alert	1	2	3	4	5	6
If I feel tired during the evening, taking a shower will liven me up	1	2	3	4	5	6
I feel tired when I have a cold	1	2	3	4	5	6
After exercise I feel invigorated	1	2	3	4	5	6
I feel at my most alert later in the day	1	2	3	4	5	6
I find it difficult to get out of bed in the morning	1	2	3	4	5	6
Cold weather makes me feel refreshed	1	2	3	4	5	6

Thank you for your co-operation
Fiona Earle
Department of Psychology
University of Hull

Appendix 2: Initial 6 Factor Solution

	Component					
ITEMS	**Gen**	**Phys**	**Ment**	**Morn**	**Even**	**Naps**
enjoy hard physical work		-.635				
get tired for no reason					.465	
too tired to think for long periods			.359			
feel energetic over the whole day				-.458		
find physical activity hard		.543				
resist external distraction/conc			-.584			
listening to others problems	.332					
sleepy at work in the afternoon					.345	.383
winter weather / depressed and weary	.360					
good night sleep feel revived				-.508		
quick nap perks me up						.653
difficulty alert / long journey	.357					
have cold / feel lethargic	.363					
missing meals feel tired	.486					
I wake up feeling tired					-.552	
emotional conflicts / drained	.409					
if tired prefer less demanding	.309					
get worn out by prolonged mental activity	.432		.388			
expect to feel lively		-.393			-.382	
working without break makes me feel weary	.511					
not an energetic person		.628		.413		
worn out after exerting myself physically		.559				
trouble sleeping at night				.467		
tired when need to concentrate hard for long periods	.498		.350			
being physically active makes me mentally alert		-.624				
no trouble remaining alert during the day			-.418			
get tired when its really hot	.425					
don't feel like making an effort when Im tired	.443					
If I stay in for the evening I feel sleepy	.441					
I get up early in the morning				-.640		
I am regarded by my friends as having bags of energy		-.515				
My capacity for concentration falls off later in the day	.377				.451	
I like to take naps during the day						.781
sustain a high level of mental effort for long periods			-.748			
If I am doing something when I am tired I tend to skip less important details	.389					

I take a long time to get going in the morning				.612		
I am able to concentrate for long periods without lapse of attention			-.803			
I only feel alert if I eat at the right times	.431					
Having a hot bath makes me feel sleepy	.355					
My current sleep pattern works well for me				-.433		
Strenuous physical work exhausts me	.398	.585				
I come across to others as being a bit lethargic		.386		.313		
tired dealing with problems	.501					
great capacity for dealing with emotional problems						
I feel tired at the end of a full working day	.555					
I can sustain attention even on quite monotonous tasks			-.588			
I am full of energy first thing in the morning				-.637	-.320	
physical work leaves me feeling worn out	.417	.622				
When Im hungry I feel drained of energy	.496					
After waking I feel sleepy for some time				.525		
I get tired when I have been trying to deal with a stressful problem	.573					
When I am very tired I am prepared to cut a few corners	.329		.310	.304		
I find other peoples problems quite wearing	.385		.455			
I have trouble falling asleep at night					-.379	
lose interest if not mentally stimulated	.322					
low overall vitality		.486			.319	
refreshed after sleep					-.746	
go to bed early				-.531		
get bored if not much of interest	.327					
prolonged mental activity wears me out	.554		.314			
wake up before need to get up				-.415		
feel sleepy after alcohol					.307	
enjoy being physically active		-.710				
sluggish after long train journeys	.432					
sport and ex feel tired	.305	.529				
naps good way to keep fresh						.795
eating big meals makes me sleepy	.313					
tired in evenings after work	.586					
work under pressure / worn out	.554					
driving several hours / feel tired	.438					
time with friends is invigorating						
prefer difficult work in afternoons				.331		
like a nap in middle of day						.788
tend to feel sleepy in afternoons	.350					.551

recover quickly from physical work		-.665				
feel lively in mornings				-.601	-.331	
repetitive work / feel weary	.393					
when get up early feel tired all day				.465		
feel more lively when out of doors	.304	-.308				
happy to slump in front of tv for evening						
feel alert all day			-.421		-.384	
feel more tired after over sleeping						
doing nothing makes me feel weary	.380					
left to myself / drop off in evenings						.513
feel rested after long train journeys						
stay up late to work / feel terrible next day	.392					
lose sleep / tend to drop off during day			.343			.358
cup of coffee makes me feel more alert						
feel most energetic around middle of the day						
like the feeling from hard physical work		-.710				
can stay up without feeling any ill effect				.306		
few drinks / lively and alert				.390		
feel tired in evening / shower will liven me up						
feel tired when I have a cold	.430					
after exercise feel invigorated		-.679				
feel most alert later in the day				.493		
difficult to get out of bed in the mornings				.659		
cold weather makes me feel refreshed						

Appendix 3: Trait Scale (2)

GENERAL TIREDNESS QUESTIONNAIRE 2

Name **Sex** **Age**
Occupation

The following list of statements has been compiled from comments people have made about their general feelings of tiredness and energy in different circumstances. Please read each one carefully and indicate how much it applies to your normal everyday state – that is, **how you are generally***. Circle one of the numbers on each line: **1** = disagree strongly; **2** = disagree; **3** = neither agree nor disagree; **4** = agree; **5** = agree strongly.*
- i.e. the higher the number the more you agree that this statement applies to the way you generally feel.

Note: Although some of the statements are similar to others please respond to each one separately. Go through the list from beginning to end without leaving gaps or going back to change responses. Of course, there are no right or wrong answers – we are interested only in measuring the variety of experiences that people have in relation to feelings of tiredness and energy across a wide range of situations. A small number of items may not apply to everyone. If an item does not apply to you at all, please write N/A next to it

Thinking of how I am generally…..	***Disagree***				***Agree***
1) I get tired for no reason	1	2	3	4	5
2) I wake up feeling tired	1	2	3	4	5
3) My capacity for concentration falls off later in the day	1	2	3	4	5
4) I feel tired when I do a lot of physically demanding tasks	1	2	3	4	5
5) I feel tired at the end of a full working day	1	2	3	4	5
6) I feel alert most of the time	1	2	3	4	5
7) I get tired when I need to concentrate for long periods	1	2	3	4	5
8)I take a long time to get going in the morning	1	2	3	4	5
9) I feel worn out after exerting myself physically	1	2	3	4	5
10) Prolonged mental activity wears me out	1	2	3	4	5
11) I feel tired in the evenings after working all day	1	2	3	4	5
12) I am not a lively person	1	2	3	4	5
13) Strenuous physical work exhausts me	1	2	3	4	5
14) However tired I am I feel refreshed after sleep	1	2	3	4	5
15) I don't feel like making much of an effort when I am tired	1	2	3	4	5
16) When I am tired I tend to cut a few corners to get things done	1	2	3	4	5
17) I think of myself as having bags of energy	1	2	3	4	5
18) If I am doing something when tired I tend to skip less important details	1	2	3	4	5
19) Physical work leaves me feeling worn out	1	2	3	4	5
20) When I have to get up early I feel tired all day	1	2	3	4	5
21) I feel most alert later in the day	1	2	3	4	5
22) I am unable to sustain a high level of mental effort for long periods	1	2	3	4	5
23) I am able to concentrate for long periods without lapses of attention	1	2	3	4	5
24) If I am feeling tired during work I try to do less demanding tasks	1	2	3	4	5

Appendix 4: Second 6 Factor Solution

	Component					
ITEMS	**Phys**	**Gen**	**Ment**	**Even**	**Strat**	**Morn**
get tired for no reason		.634				
I wake up feeling tired		.662				
if tired prefer less demanding			.424		.365	
not a lively person	.470	.626				
worn out after exerting myself physically	.708					
tired when need to concentrate hard for long periods			.618			
don't feel like making an effort when I'm tired					.557	
I think of myself as having bags of energy	-.310	-.629				
My capacity for concentration falls off later in the day				.373		
unable to sustain a high level of mental effort for long periods			.817			
If I am dooing something when I am tired I tend to skip less important details					.761	
I take a long time to get going in the morning						.490
I am able to concentrate for long periods without lapse of attention			-.766			
Strenuous physical work exhausts me	.752					
I feel tired at the end of a full working day				.724		
physical work leaves me feeling worn out	.778					
When I am tired I tend to cut a few corners					.787	
refreshed after sleep		-.662				
prolonged mental activity wears me out			.549	.365		
tired in evenings after work				.784		
when get up early feel tired all day						.321
feel alert most of the time		-.571	-.370			
feel most alert later in the day						.787
I feel tired when I do a lot of physically demanding tasks	.799					

Appendix 5: Composite Validation Questionnaire

GENERAL FEELINGS QUESTIONNAIRE

On the following pages, there are five questionnaires. Each has a different focus, and some have different ways of responding, so it is important that you read the specific instructions at the top of each page. Each questionnaire asks you about how you generally feel, and we are interested in finding out how you respond to the different types of questions.

It is important that you think about each question separately. Although some of the questions may appear to be asking you very similar things, this is not designed to catch you out, and you should not try to answer questions in a way which fits in with previous answers. It does not matter if some of your answers seem a little contradictory. Just answer each question as honestly as possible and do not spend too long on any one item. There is no right or wrong answer, we are only interested in measuring the variety of experiences that people have across a wide range of situations.

To help us interpret the information you provide, is important that we know a little bit about you and your life. Please complete the section below, before responding to the following questionnaires.

All information you provide is strictly confidential.

About You:

1. Age: ____________________

2. Gender: ______________________________

3. Marital status: ____________

4. Do you have any children: YES / NO

5. If YES, how many children do you have and what age/s are they? _________________________

About your Job:

6. Do you work: full-time / part-time / flexi-time

7. Does your job involve working at shifts or unconventional hours? (Please specify the nature of your working hours): ____

8. To what extent does your job make the following **demands** on you? Please rate each type of demand on the following scale:

1 = not at all; 2 = a little; 3 = some; 4 = quite a bit; 5 = a great deal.

	Not at all				A great deal
a. Mental demands	1	2	3	4	5
b. Physical demands	1	2	3	4	5
c. Emotional demands	1	2	3	4	5

9. How well do you feel you generally **cope** with these demands? Please circle below

Not at all well (1)	It varies (2)	Quite well (3)	Very well (4)

10. To what extent does your job offer the following opportunities?

	Not at all				A great deal
a. Personal control Opportunity to decide how, when and in which order jobs should be carried out	1	2	3	4	5
b. Personal support Help and support from others	1	2	3	4	5

PLEASE CONTINUE

1. GENERAL TIREDNESS QUESTIONNAIRE

The first set of questions has been compiled from comments people have made about their general feelings of tiredness and energy in different circumstances. Please read each one carefully and indicate how much it applies to your normal everyday state – that is, **how you are generally.**

Circle one of the numbers on each line:

1 = disagree strongly; 2 = disagree; 3 = neither agree nor disagree; 4 = agree; 5 = agree strongly. i.e. the higher the number the more you agree that this statement applies to the way you generally feel.

Thinking of how I feel generally	Disagree				Agree
1) I get tired for no reason	1	2	3	4	5
2) I feel tired when I do a lot of physically demanding tasks	1	2	3	4	5
3) My capacity for concentration falls off later in the day	1	2	3	4	5
4) I feel alert most of the time	1	2	3	4	5
5) I feel tired at the end of a full working day	1	2	3	4	5
6) I take a long time to get going in the morning	1	2	3	4	5
7) I get tired when I need to concentrate for long periods	1	2	3	4	5
8) Prolonged mental activity wears me out	1	2	3	4	5
9) I feel worn out after exerting myself physically	1	2	3	4	5
10) I am not a lively person	1	2	3	4	5
11) I feel tired in the evenings after working all day	1	2	3	4	5
12) When I am tired I tend to cut a few corners to get things done	1	2	3	4	5
13) Strenuous physical work exhausts me	1	2	3	4	5
14) If I am doing something when tired I tend to skip less important	1	2	3	4	5
15) I don't feel like making much of an effort when I am tired	1	2	3	4	5
16) When I have to get up early I feel tired all day	1	2	3	4	5
17) I think of myself as having bags of energy	1	2	3	4	5
18) I am unable to sustain a high level of mental effort for long periods	1	2	3	4	5
19) Physical work leaves me feeling worn out	1	2	3	4	5
20) If I am feeling tired during work I try to do less demanding tasks	1	2	3	4	5
21) I feel most alert later in the day	1	2	3	4	5
22) I am able to concentrate for long periods without lapses of attention	1	2	3	4	5

PLEASE CONTINUE

2. HARDINESS QUESTIONNAIRE

As before, please indicate your response to the following items by circling the number which best describes how you feel generally. **The higher the number, the more you agree with the statement.**

	Disagree			Agree	
1) I usually find something to motivate me	1	2	3	4	5
2) I generally feel in control	1	2	3	4	5
3) I generally feel that I am a worthwhile person	1	2	3	4	5
4) Challenges usually bring out the best in me	1	2	3	4	5
5) When working with other people I am usually quite influential	1	2	3	4	5
6) Unexpected changes to my schedule generally throw me	1	2	3	4	5
7) I don't usually give up under pressure	1	2	3	4	5
8) I am generally confident in my own abilities	1	2	3	4	5
9) I usually find myself just going through the motions	1	2	3	4	5
10) At times I expect things to go wrong	1	2	3	4	5
11) "I just don't know where to begin" is a feeling I usually have when presented with several things to do at once	1	2	3	4	5
12) I generally feel that I am in control of what happens in my life	1	2	3	4	5
13) However bad things are, I usually feel they will work out positively in the	1	2	3	4	5
14) I often wish my life was more predictable	1	2	3	4	5
15) Whenever I try to plan something, unforeseen factors usually seem to	1	2	3	4	5
16) I generally look on the bright side of life	1	2	3	4	5
17) I usually speak my mind when I have something to say	1	2	3	4	5
18) At times I feel completely useless	1	2	3	4	5
19) I can generally be relied upon to complete the tasks I am given	1	2	3	4	5
20) I usually take charge of a situation when I feel it is appropriate	1	2	3	4	5
21) I generally find it hard to relax	1	2	3	4	5
22) I am easily distracted from tasks that I am involved with	1	2	3	4	5
23) I generally cope well with any problems that occur	1	2	3	4	5
24) I do not usually criticise myself even when things go wrong	1	2	3	4	5
25) I generally try to give 100%	1	2	3	4	5
26) When I am upset or annoyed I usually let others know	1	2	3	4	5
27) I tend to worry about things well before they actually happen	1	2	3	4	5
28) I often feel intimidated in social gatherings	1	2	3	4	5
29) When faced with difficulties I usually give up	1	2	3	4	5
30) I am generally able to react quickly when something unexpected happens	1	2	3	4	5
31) Even when under considerable pressure I usually remain calm	1	2	3	4	5
32) If something can go wrong, it usually will	1	2	3	4	5
33) Things just usually happen to me	1	2	3	4	5
34) I generally hide my emotion from others	1	2	3	4	5
35) I usually find it difficult to make a mental effort when I am tired	1	2	3	4	5
36) When I make mistakes I usually let it worry me for days after	1	2	3	4	5
37) When I am feeling tired I find it difficult to get going	1	2	3	4	5
38) I am comfortable telling people what to do	1	2	3	4	5
39) I can normally sustain high levels of mental effort for long periods	1	2	3	4	5
40) I usually look forward to changes in my routine	1	2	3	4	5
41) I feel that what I do tends to make no difference	1	2	3	4	5
42) I usually find it hard to summon enthusiasm for the tasks I have to do	1	2	3	4	5
43) If I feel somebody is wrong, I am not afraid to argue with them	1	2	3	4	5
44) I usually enjoy a challenge	1	2	3	4	5
45) I can usually control my nervousness	1	2	3	4	5
46) In discussions, I tend to back-down even when I feel strongly about	1	2	3	4	5
47) When I face setbacks I am often unable to persist with my goal	1	2	3	4	5
48) I can usually adapt myself to challenges that come my way	1	2	3	4	5

3. SELF EVALUATION QUESTIONNAIRE

The STAI questionnaire was removed from this thesis for copyright reasons.

PLEASE CONTINE

4. MULTIDIMENSIONAL FATIGUE INVENTORY

The following statements ask you **how you have been feeling __lately__**. There is for example, the statement: *" I feel relaxed"* If you think that this is true, that you have indeed been feeling relaxed lately, please place an X in the far right box; like this:

No, that is not true ☐ ☐ ☐ ☐ ☒ Yes, this is true

The more you disagree with the statement, the more you can place an X in the direction of "no, that is not true". Please do not miss out a statement and place one X next to each statement.

1. I feel fit	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
2. Physically I feel only able to do a little	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
3. I feel very active	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
4. I feel like doing all sorts of nice things	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
5. I feel tired	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
6. I think I do a lot in a day	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
7. When I am doing something I can keep my thoughts on it	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
8. Physically I can take on a lot	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
9. I dread having to do things	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
10. I think I do very little in a day	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
11. I can concentrate well	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
12. I am rested	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
13. It takes a lot of effort to concentrate on things	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
14. Physically I feel in a bad condition	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
15. I have a lot of plans	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
16. I tire easily	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
17. I get little done	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
18. I don't feel like doing anything	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
19. My thoughts easily wander	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true
20. I feel I am in excellent condition	No, that is not true	☐ ☐ ☐ ☐ ☐	Yes, this is true

GENERAL WELL BEING SCALE:

Below are a set of statements which again ask how you generally feel in terms of your well-being. Each possible response is associated with a letter. Circle the letter which best describes how you feel.

Thinking of how I feel generally.....

1. I feel tense and 'wound up':
- a. Most of the time
- b. A lot of the time
- c. From time to time, occasionally
- d. Not at all

2. I still enjoy the things I used to enjoy:
- a. Definitely as much
- b. Not quite so much
- c. Only a little
- d. Hardly at all

3. I get a sort of frightened feeling as if something awful is about to happen:
- a. Very definitely and quite badly
- b. Yes, but not too badly
- c. A little, but it doesn't worry me
- d. Not at all

4. I can laugh and see the funny side of things:
- a. As much as I always could
- b. Not quite so much now
- c. Definitely not so much now
- d. Not at all

5. Worrying thoughts go through my mind:
- a. A great deal of the time
- b. A lot of the time
- c. From time to time but not too often
- d. Only occasionally

6. I feel cheerful:
- a. Not at all
- b. Not often
- c. Sometimes
- d. Most of the time

7. I can sit at ease and feel relaxed:
- a. Definitely
- b. Usually
- c. Not often
- d. Not at all

8. I feel as if I am slowed down:
- a. Nearly all the time
- b. Very often
- c. Sometimes
- d. Not at all

9. I get a sort of feeling like 'butterflies' in the stomach:
- a. Not at all
- b. Occasionally
- c. Quite often
- d. Very often

10. I have lost interest in my appearance:
- a. Definitely
- b. I don't take as much care as I should
- c. I may not take quite so much care
- d. I take just as much care as ever

11. I feel restless as if I have to be on the move:
- a. Very much indeed
- b. Quite a lot
- c. Not very much
- d. Not at all

12. I look forward with enjoyment to things:
- a. As much as I ever did
- b. Rather less than I used to
- c. Definitely less than I used to
- d. Hardly at all

13. I get sudden feelings of panic:
- a. Very often indeed
- b. Quite often
- c. Not very often
- d. Not at all

14. I can enjoy a good book, or radio, or TV programme:
- a. Often
- b. Sometimes
- c. Not often
- d. Very seldom

THANKYOU FOR YOUR HELP

Appendix 6: Reliability Statistics for the GTQ

General Tiredness Questionnaire Sub Scales	No. of Items	Cronbach's Alpha
GTQ General Fatigue	4	0.80
GTQ Morning Tiredness	3	0.76
GTQ Evening Tiredness	3	0.81
GTQ Physical Fatigue	4	0.90
GTQ Mental Fatigue	4	0.84
GTQ Mental Strategies	4	0.82

Appendix 7: The Feelings Questionnaire (state fatigue scale)

NAME:

Feelings 1

Below are a set of statements which describe a range of feelings. Please indicate to what extent you agree with each statement - **considering how you feel right now**
1 = strongly disagree; 5 = strongly agree

	Disagree				Agree
I feel mentally tired	1	2	3	4	5
I feel bored	1	2	3	4	5
I feel somewhat sleepy	1	2	3	4	5
I feel detached / uninterested	1	2	3	4	5
I don't feel like making much of an effort	1	2	3	4	5
I feel like closing my eyes and having a nap	1	2	3	4	5
I feel worn out physically	1	2	3	4	5
I feel uneasy	1	2	3	4	5
I feel unable to concentrate	1	2	3	4	5
I feel tense / on edge	1	2	3	4	5
I feel irritated and annoyed	1	2	3	4	5
I feel wide awake	1	2	3	4	5
I feel mentally drained	1	2	3	4	5
I feel physically tired	1	2	3	4	5
I feel drowsy	1	2	3	4	5

Appendix 8: Example of Validation diary page

Friday: **Today's date / /**

Time starting work _____ time finishing work _____

not at work_____

Total hours/mins of breaks/free time ______hrs/______mins

Part 1: Current Tiredness:

Below are a set of statements which describe a range of feelings. Please indicate to what extent you agree with each statement – **considering how you feel right now**

	Strongly disagree			Strongly agree	
I feel mentally tired	1	2	3	4	5
I feel bored	1	2	3	4	5
I feel somewhat sleepy	1	2	3	4	5
I feel detached/ uninterested	1	2	3	4	5
I don't feel like making much of a effort	1	2	3	4	5
I feel like closing my eyes and having a nap	1	2	3	4	5
I feel worn out physically	1	2	3	4	5
I feel uneasy	1	2	3	4	5
I feel unable to concentrate	1	2	3	4	5
I feel tense/ on edge	1	2	3	4	5
I feel irritated and annoyed	1	2	3	4	5
I feel wide awake	1	2	3	4	5
I feel mentally drained	1	2	3	4	5
I feel physically tired	1	2	3	4	5
I feel drowsy	1	2	3	4	5

Total number of hours slept: _______hrs

Write the number of the statement that best describes how you feel right now:

____

1 = Fully alert, wide awake, very peppy

2 = Very lively, responsive, not at peak

3 = Okay, somewhat fresh

4 = A little tired, less than fresh

5 = Moderately let down

6 = Extremely tired

7 = Completely exhausted, unable to function

Part 2: Present Mood:

Please indicate how you have felt today in terms of the following dimensions, based on different aspects of moods. Circle the number that best describes how you feel, between each pair of extremes.

Enthusiastic	1	2	3	4	5	6	7	8	9	Miserable
Weary	1	2	3	4	5	6	7	8	9	Lively
Relaxed	1	2	3	4	5	6	7	8	9	Tense
Depressed	1	2	3	4	5	6	7	8	9	Optimistic
Energetic	1	2	3	4	5	6	7	8	9	Tired
On edge	1	2	3	4	5	6	7	8	9	At ease

Part 3: Demands and opportunities:

Please indicate the level of each kind of demand made on you today, and also the level of control and support experienced.

	Low								High
Mental demands	1	2	3	4	5	6	7	8	9
Emotional demands	1	2	3	4	5	6	7	8	9
Physical demands	1	2	3	4	5	6	7	8	9
Personal control	1	2	3	4	5	6	7	8	9
Personal support	1	2	3	4	5	6	7	8	9

Please feel free to add any further comments:

Appendix 9: Within person correlations between daily demands and FQ fatigue scales

Individual 1 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	.31				
FQ sleep fatigue	**.73**	.16			
mental demands	.24	.36	.33		
emotional demands	.20	.46	.25	**.95**	
physical demands	.39	-.28	.19	-.15	-.01

Individual 2 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	-.10				
FQ sleep fatigue	**.73**	-.08			
mental demands	.21	-.25	-.048		
emotional demands	.33	.10	-.001	**.81**	
physical demands	.00	**.83**	-.082	-.43	-.03

Individual 3 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	.33				
FQ sleep fatigue	**.70**	.33			
mental demands	-.23	-.18	-.32		
emotional demands	.15	.10	-.15	.48	
physical demands	.17	.04	.32	**-.62**	-.26

Individual 4 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	**.74**				
FQ sleep fatigue	**.75**	.45			
mental demands	**.47**	.40	.20		
emotional demands	.52	**.50**	.33	.50	
physical demands	.28	.29	.00	.25	**.78**

Individual 5 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	**-.72**				
FQ sleep fatigue	**.81**	**-.68**			
mental demands	**.86**	**-.68**	**.63**		
emotional demands	**.87**	**-.68**	**.67**	**.96**	
physical demands	**-.75**	**.89**	**-.76**	**-.73**	**-.70**

Individual 6 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	**.86**				
FQ sleep fatigue	**.80**	**.71**			
mental demands	.22	-.01	.20		
emotional demands	.21	-.06	.05	**.67**	
physical demands	-.08	.27	-.01	**-.57**	-.52

Individual 7 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	-.15				
FQ sleep fatigue	.56	.47			
mental demands	.36	.10	**.92**		
emotional demands	.40	-.66	-.02	.40	
physical demands	-.35	**.91**	.47	-.62	.29

Individual 8 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	.52				
FQ sleep fatigue	**.76**	.34			
mental demands	**.88**	.37	**.88**		
emotional demands	**.55**	.23	**.58**	**.69**	
physical demands	.44	.50	.45	.50	.15

Individual 9 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	.36				
FQ sleep fatigue	.69	.39			
mental demands	.37	-.34	.68		
emotional demands	-.09	.24	.50	.46	
physical demands	-.57	-.66	-.21	.40	.37

Individual 10 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	.38				
FQ sleep fatigue	**.72**	**.65**			
mental demands	**.76**	.16	.41		
emotional demands	.00	-.24	-.49	.37	
physical demands	-.08	-.28	-.41	.34	**.80**

Individual 11 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	26				
FQ sleep fatigue	.17	.10			
mental demands	.35	.16	.03		
emotional demands	-.10	-.22	.23	.36	
physical demands	-.28	.40	-.47	.35	-.03

Individual 12 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	**.65**				
FQ sleep fatigue	**.78**	.70			
mental demands	**.66**	**.64**	.41		
emotional demands	**.61**	**.63**	.42	**.82**	
physical demands	.26	.30	.02	**.62**	.52

Individual 13 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	**.72**				
FQ sleep fatigue	**.63**	**.62**			
mental demands	.29	.50	.16		
emotional demands	.34	.52	.36	**.69**	
physical demands	**.58**	.51	.27	.33	**.64**

Individual 14 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	.53				
FQ sleep fatigue	**.97**	.42			
mental demands	**.87**	**.83**	.78		
emotional demands	**.94**	.47	**.97**	**.83**	
physical demands	.78	**.82**	.65	**.96**	.43

Individual 15 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	-.08				
FQ sleep fatigue	**.53**	.20			
mental demands	**.73**	-.19	.41		
emotional demands	.33	-.15	-.09	.51	
physical demands	.31	.29	-.21	.47	**.64**

Individual 16 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	**.60**				
FQ sleep fatigue	.17	**.58**			
mental demands	-.19	.34	.29		
emotional demands	.48	.00	-.09	-.24	
physical demands	.02	-.04	-.16	.26	.03

Individual 17 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	.28				
FQ sleep fatigue	**.55**	.17			
mental demands	.06	-.38	-.44		
emotional demands	.16	-.07	**-.54**	**.72**	
physical demands	-.45	.43	-.33	-.07	-.10

Individual 18 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	-.19				
FQ sleep fatigue	**.81**	-.13			
mental demands	.23	-.08	.10		
emotional demands	.24	.51	.02	**.59**	
physical demands	.26	-.06	.27	**.69**	.35

Individual 19 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	**.85**				
FQ sleep fatigue	**.79**	**.93**			
mental demands	**.53**	.62	.31		
emotional demands	.57	.66	.43	**.89**	
physical demands	.12	.10	-.07	.42	.03

Individual 20 – Correlations between daily demands and FQ fatigue scales

	FQ mental fatigue	FQ physical fatigue	FQ sleep fatigue	mental demands	emotional demands
FQ –physical fatigue	.32				
FQ sleep fatigue	**.79**	**.59**			
mental demands	**.64**	-.16	.44		
emotional demands	.04	.12	.03	.41	
physical demands	**-.55**	.22	-.35	**-.57**	.33

Appendix 10: CAMS System Knowledge Questionnaire

SYSTEM KNOWLEDGE QUESTIONNAIRE

1. What functions do each of the following have in the system?

a. Heater

____

____

b. Vent

____

____

c. Carbon Dioxide Scrubber

____

____

d. Dehumidifier

____

____

2. Under normal conditions, the values of the sub-systems rise and fall in regular patterns. What causes this?

____

____

____

3. List all the events in the system which could reduce PRESSURE

____

____

____

____

Please turn over

4. Mark the following statements True “T” or False “F”

a. The nitrogen valve opens when pressure falls because the automatic controller responds to the fall in pressure ________________

b. The mixer unit operates when activated by the control system ________________

c. Oxygen flow is reduced when Nitrogen is too high ________________

d. The cooler is automatically activated by the control system when the temperature rises above a set point ________________

5. The following questions ask about your knowledge of relationships between the different system components.

Example 1.

	Increase	Decrease	Little no effect
What effect will the Heater have on cabin temperature?	☑	☐	☐

Please explain why: *The heater has a direct effect on cabin temperature*

	Increase	Decrease	Little / no effect
a. What effect does the CO_2 scrubber have on cabin air pressure?	☐	☐	☐

Please explain why

____

____

____

	Increase	Decrease	Little / no effect
b. What effect does the heater have on humidity levels?	☐	☐	☐

Please explain why

____

____

____

	Increase	Decrease	Little / no effect
c. What effect does the dehumidifier have on O_2 levels?	☐	☐	☐

Please explain why

____

____

____

Appendix 11: Experimental Schedule (Exp. 1)

SEQUENCE FOR GROUP 1	SEQUENCE FOR GROUP 2
TIMING (minutes)	TIMING (minutes)
When 5 (3, 4.5, 6, 7.5, 9)	When 5 (3, 4.5, 6, 7.5, 9)
When 4 (3, 5, 7, 9)	When 4 (3, 5, 7, 9)
When 3 (3, 6, 9)	When 3 (3, 6, 9)
When 2 (3, 9)	When 2 (3, 9)
When 1 (3)	When 1 (3)
ORDER	ORDER
A:	**A:**
H (180) O2 (540)	CO2 (180) H (300) O2 (420) T (540)
T (180)	O2 (180) P (270) T (360) CO2 (450) H (540)
O2 (180) P (270) T (360) CO2 (450) H (540)	CO2 (180)
CO2 (180) P (540)	T (180) P (270) CO2 (360) H (450) O2 (540)
O2 (180)	H (180) P (300) O2 (420) CO2 (540)
CO2 (180) P (360) T (540)	T (180) P (540)
H (180) P (300) O2 (420) T (540)	O2 (180) P (360) H (540)
B:	**B:**
H (180)	T (180) P (270) CO2 (360) H (450) O2 (540)
T (180) P (270) CO2 (360) H (450) O2 (540)	H (180) CO2 (540)
H (180) P (540)	T (180) P (300) O2 (420) H (540)
O2 (180) T (300) CO2 (420) H (540)	P (180)
P (180)	O2 (180) P (270) T (360) CO2 (450) H (540)
T (180) CO2 (540)	T (180) P (360) CO2 (540)
T (180) P (360) O2 (540)	CO2 (180) H (300) O2 (420) T (540)
C:	**C:**
O2 (180) T(540)	CO2 (180) P (360) T (540)
P (180)	O2 (180) P (300) T (420) CO2 (540)
CO2 (180) P (300) T (420) O2 (540)	H (180)
O2 (180) P (270) T (360) CO2 (450) H (540)	O2 (180) P(540)
O2 (180) H (360) CO2 (540)	O2 (180) P (270) T (360) CO2 (450) H (540)
CO2 (180)	T (180) CO2 (300) O2 (420) H (540)
H (180) P (540)	T (180) P (270) CO2 (300) H (450) O2 (540)

Appendix 12: Workload Assessment Questionnaire

POST TASK QUESTIONNAIRE

Name ____________________

Considering the work you have completed……

1. How much **attention** did the task require?
Very little _ A Great Deal

2. To what extent did the task make **physical** demands on you?
Very little _ A Great Deal

3. How much did the task require the operator to **solve difficult problems**?
Very little _ A Great Deal

4. How much did the task require the operator to **intervene** or change the state of the system?
Very little _ A Great Deal

5. How much did the task allow the operator to decide **what to do**, or **how** to do it?
Very little _ A Great Deal

6. How much did the task allow the operator to decide **when** to carry out an action, or take a break?
Very little _ A Great Deal

7. How much would it **matter** if you did not perform as well as possible?
Very little _ A Great Deal

8. How much **effort** did you put into the task?
Very little _ A Great Deal

9. How much **time pressure** did you experience?
Very little _ A Great Deal

10. How much **frustration** did you experience?
Very little _ A Great Deal

11. How well did your **performance** meet the overall goals of maintaining the system without error?
Not at all well _ Very well

Appendix 13: Two Effort Questionnaires

ENGAGEMENT WITH GOALS

How concerned were you with keeping to the following goals:

1. The primary task of maintained the five systems within the normal operating range?

Not at all _ Very much

2. The secondary task of acknowledging alarms?

Not at all _ Very much

3. The secondary task of carrying out system checks on time?

Not at all _ Very much

EFFORT ALLOCATION

If you consider that the total amount of effort you put into the task consisted of 100 units of effort, how did you distribute them across the 3 tasks?

Divide your 100 units into 3 to represent the way you allocated your effort.

Maintaining the system within the normal operating range ☐

Acknowledging alarms as quickly as possible ☐

Carrying out system checks at the appropriate time ☐