UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

A SYSTEMS APPROACH TO THE ASSESSMENT OF MENTAL WORKLOAD IN A SAFETY-CRITICAL ENVIRONMENT

by

Adele Krüger

Submitted in partial fulfilment of the requirements for the degree

Philosophiae Doctor (Bio-Systems)

in the

Faculty of Engineering, Built Environment and Information Technology

UNIVERSITY OF PRETORIA

PRETORIA

May 2008

SUMMARY

TITLE:	A systems approach to the assessment of mental workload in a safety-critical environment
CANDIDATE:	Adele Krüger
PROMOTOR:	Dr. P.J. Cilliers
CO-PROMOTOR:	Prof. J. A. van Tonder
DEPARTMENT:	Electrical, Electronic, and Computer Engineering
DEGREE:	Philosophiae Doctor (Bio-Systems)

The objective of this study is to develop a quantified method for determining the mental workload imposed on train control officers (TCOs) and to express this mental workload by means of an index that is objective and can stand up to the tests of validity and reliability. The method addresses an existing operational shortcoming in Spoornet train control operations and could be used as a tool for predicting the mental workload imposed on operators at particular train control centres. The method could be applied to manage and improve operational safety in the rail transport environment. A participative systems approach was followed in the development of the measuring methodology. A work group comprising expert users of the specific train control system was involved in identifying task factors and assigning weights for task and moderating factors. The newly developed Mental Workload Index (MWLI) consists of three task factors and eleven moderating factors, each with a different weight in terms of its contribution to overall mental workload. The work group performed several iterations to reach final consensus on the following task factors and their respective contributions to the MWLI: the number of data transactions, the number of authorisations, and the number of communications via telephone and radio. The systems approach used in the development process is discussed, and the final index with the task and moderating factors is presented. In conclusion, the value and possible application of the MWLI are discussed. The MWLI is shown to provide an objective method for the assessment and prediction of mental workload in the train control environment.

Keywords: *Mental workload, Mental workload assessment, Mental workload prediction, Workload, Stress, Cognitive load.*

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

OPSOMMING

TITEL: A systems approach to the assessment of mental workload in a safety critical environment

KANDIDAAT: Adele Krüger

PROMOTOR: Dr. P.J. Cilliers

MEDE-PROMOTOR: Prof. J.A. van Tonder

DEPARTEMENT: Elektries-, Elektronies- en Rekenaaringenieurswese

GRAAD: Philosophiae Doctor (Bio-Sisteme)

Die doel van die studie is die ontwikkeling van ’n gekwantifiseerde metode om die kognitiewe werkslas van treinbeheeramptenare (TBAs) te bepaal en om hierdie kognitiewe werkslas uit te druk as ’n indeks wat objektief en kwantifiseerbaar is en wat die toetse van geldigheid en betroubaarheid kan deurstaan. Hierdie metode spreek ’n bestaande operasionele tekortkoming aan en kan gebruik word as ’n instrument om die kognitiewe werkslas wat op ’n operateur geplaas word by ’n spesifieke treinbeheersentrum, te voorspel. Die metode kan gebruik word om operasionele veiligheid in die spoorverkeeromgewing te bestuur en verbeter. ’n Deelnemende sisteembenadering is gevolg in die ontwikkelingsproses. ’n Werkgroep, bestaande uit ekspert gebruikers van die spesifieke treinbeheerstelsel was betrokke by die identifisering van taak- en modererende faktore sowel as by die toekenning van gewigte aan hierdie faktore. Die nuut-ontwikkelde Mental Workload Index (MWLI) bestaan uit drie taakfaktore en elf modererende faktore, met verskillende gewigte na aanleiding van elkeen se bydrae tot die totale werkslas. Die werkgroep het verskillende iterasies uitgevoer ten einde konsensus te verkry oor die volgende taakfaktore en hul onderskeie bydrae tot MWLI: aantal datatransaksies, aantal magtigings en aantal kommunikasies oor die telefoon en radio. Die sisteembenadering wat gevolg is in die ontwikkelingsproses word bespreek en die finale indeks met die taak- en modererende faktore word aangebied. Ten slotte word die waarde en moontlike toepassing van die MWLI bespreek. Die MWLI word aangebied as ’n objektiewe metode vir die meting en voorspelling van kognitiewe werkslas in die treinbeheer omgewing.

Sleutelwoorde: *Mental workload, Mental workload assessment, Mental workload prediction, Workload, Stress; Cognitive load.*

Dedicated to my mother – a courageous woman

ACKNOWLEDGEMENTS

I wish to thank the following individuals who contributed to this study by generously sharing and giving their knowledge, experience, expertise, support and encouragement throughout the process:

- Dr. Pierre Cilliers, my study leader, for his continuous support, positive energy and constructive criticism.
- Prof. Jan van Tonder, the co-study leader, for being a source of help and support.
- Johann, for his unconditional love and understanding, tolerance, support and unselfish sacrifices.
- My mother, for her unwavering belief in me.
- Herman Lemmer and the Spoornet expert work group, for their participation and assistance in the development phase of the Mental Workload Index.
- The Spoornet Operational and Section Managers, for their support with arrangements in the validation process at the respective train control centres.
- The Train Control Officers, for their enthusiastic participation in the study.
- Prof. Retha Viljoen, Dr. Nico Claassen and Lynton Hazelhurst of the Physiology Department, University of Pretoria, for their guidance and hard work on the pilot study and the validation of the Mental Workload Index.
- Thys de Vries, operational researcher at the Defence Research Centre for facilitating the sessions with the workgroup.
- Dr. Piet Becker of the Medical Research Council for his expert input in the statistical analyses.
- Russell de la Porte, for his thorough and meticulous editing of this document.

LIST OF ABBREVIATIONS

AFFTC	Air Force Flight Test Centre
ATC	Air Traffic Control
BMI	Body Mass Index
CNS	Central Nervous System
CTC	Centralised Traffic Control
DBP	Diastolic Blood Pressure
ECG	Electrocardiogram
EEG	Electroencephalogram
ERP	Event-related Potentials
ETA	Estimated Time of Arrival
ETD	Estimated Time of Departure
FRA	Federal Railroad Administration
HPA	Hypothalamo-pituitary-adrenocortical
HRV	Heart Rate Variability
IWS	Integrated Workload Scale
MWLI	Mental Workload Index
NASA-TLX	National Aeronautical and Space Administration Task Load Index
ODEC	Operational Demand Checklist
PNS	Peripheral Nervous System
RTO	Radio Train Order
SA	Situation Awareness
SAM	Sympatho-adrenomedullary
SBP	Systolic Blood Pressure
SIMS	Spoornet Information Management System
STSS	Short-term Sensory Store
SWAT	Subjective Workload Assessment Technique
TCO	Train Control Officer
TLA	Timeline Analysis
TLI	Traffic Load Index
TRB	Transportation Research Board
TWS	Track Warrant System

TABLE OF CONTENTS

PAGE

CHAPTER 1: INTRODUCTION

CHAPTER 2: LITERATURE STUDY

CHAPTER 3: METHODOLOGY

CHAPTER 4: RESULTS

5. DISCUSSION

CHAPTER 6: CONCLUSIONS

APPENDICES

CHAPTER 1

INTRODUCTION

In this chapter the background to the study is provided and the existing research gap is defined. The justification and aims of the study is provided and a scientific hypothesis is postulated.

1.1 PROBLEM STATEMENT

The South African Rail Operator, Spoornet, was faced with a potential problem when, for safety and operational reasons, it decided to effect changes in its train control systems. The problem related to the capacity of operators to authorise trains safely and efficiently by means of a train control system traditionally used in areas of low traffic volume.

Spoornet operations currently use several train control systems, of which the colour light signalling system is used in areas of high traffic volume. This system has been in use in South Africa for over five decades. In recent years there has been a steep increase in the vandalism of lights and the theft of light and points cables on Spoornet property and infrastructure. This has rendered the colour light signalling system vulnerable and has posed a threat to safe rail operations. The risk to railway operations has therefore necessitated a re-evaluation of train control systems in certain areas and the consideration of a train control system that is less vulnerable.

Another factor that has increased the need to change to a different train control system is the significant decrease in traffic volumes in certain sectors and geographical areas. Colour light signalling is normally used in areas with high train traffic volumes. The system is expensive to maintain and lower traffic volumes do not require such a sophisticated train control system. The high cost of maintaining the colour light signals in view of the low traffic volumes rendered the system no longer viable in specific areas.

The Radio Train Order (RTO) train control system is an alternative system that is currently in use, and considering safety and economy, it was a logical choice for replacing the more vulnerable and expensive colour light signalling system. It was also considered the most suitable system because the Track Warrant System (TWS), a system developed and used in North America and one that closely resembles RTO, was at the time under consideration for implementation in South Africa.

However, changing to a different train control system is not a simple matter. There are operational factors that need to be considered before implementing a specific system, such as traffic density, the nature of the train service (passenger, goods and suburban trains), geographical factors (e.g., suburban vs. rural settings), and the availability and affordability of the requisite technology. There is also a need to consider the human operators who will be affected by the changes. This study was initiated in order to investigate and provide insight and guidelines as to how the human operator will cope with the workload potentially imposed by a change from colour light signals to radio train order.

1.2 BACKGROUND

A proliferation of control and display technology in modern systems can impose heavy demands on operators' information processing capabilities. Such technology often requires the rapid sampling and integration of a large amount of information. The resultant demands could approach or exceed the limited information-processing capacities of the operators. Consequently, the need to assess the load imposed on operators' processing capacities is particularly critical in high technology systems (Eggemeier, 1988).

Xie and Salvendy (2000) make the following comments about the sensitivity of multitask environments:

> It should be noted that handling of complex systems, such as flying an aircraft, monitoring a process-control facility and interacting with computing devices can often impose multiple, concurrent task demands on operators. Safe and efficient operation of complex systems requires that information-processing demands imposed on operators do not exceed their capabilities. Processing demands that exceed these

> limits can lead to degradation in both operator and system performance. Assessment and prediction of the workload associated with multitask environments is, therefore, an important issue.

When assessing the control of complex systems by human operators it is essential to have some predictive model that will map system demands onto operators' capacities and determine the extent to which these demands exceed capacities. The explosion of modern computer and related engineering technology has led to a significant increase in the level of complexity, sophistication and degree of integration of operational train control systems. The role of the human operator has, in parallel, changed from what it once was. Many of the tasks that operators previously performed manually are now automated. The operator's role is increasingly changing to that of a system monitor, information manager and decision maker.

These developments are relevant to the Spoornet train control environment and emphasise the necessity for a tool that would allow for the assessment and prediction of mental workload in this multitask environment.

Train Control Officers (TCOs), also known as railroad or train dispatchers, are responsible for authorising the movements of trains within the context of a particular train control system on the South African Rail network. TCOs are trained in the specific train control system of the area where they work. For every train control system there is a predetermined methodology to be followed. Without an authorisation from a TCO, whether it is a green signal aspect (the railway term referring to the colour of a signal), an authorisation number, or a token or written authority, the train is not permitted to move. Thus, TCOs bear a high level of responsibility for safe train operations, with much similarity between their tasks and those of air traffic controllers.

Train dispatchers are critical control elements in the total train control system. They are responsible for both the efficiency and the safety of the operation. Their job is demanding; they are under almost constant pressure, regularly assessing complex combinations of information, making critical decisions and responding to unexpected complications. Their work environment is often far from ideal – uncomfortable, noisy and confusing (Devoe, 1974). Even thirty years after this description, the role of the TCO remains relatively unchanged.

Between 1999 and 2004 several accidents occurred in which the workload of the TCO was cited as a causal factor. In one of these accidents there was loss of life, and in the other accidents damages in the order of millions of rand were incurred. The damage to the reputation of Spoornet as a result of negative publicity was incalculable. These occurrences as well as changes in traffic volumes, and the theft and vandalism of signalling equipment have necessitated a reconsideration of the type of train control systems that might be utilised in future.

In the design of complex systems, a question that has been posed is whether the operator can perform the tasks required by the system in both normal and contingency modes. Human operator workload is a term used for describing the synthesis of task performance in both modes. One of the objectives of workload-related evaluation is therefore to confirm that the designated personnel can effectively operate, control and maintain the system (Cilliers, 1992).

Estimates of workload can determine whether specified functions and tasks allocated to human operators are feasible in terms of time and capability requirements. If demands exceed capabilities, performance may be severely compromised. If demands do not exceed capabilities, provision should be made for a sufficient margin of residual capacity or resources so that unexpected failures or environmental events may be handled to a satisfactory level (Wickens, 1992).

The issues raised by the two authors above – the ability to operate the system and the feasibility of task allocation in terms of operators' capabilities – are the gaps that this study aims to address.

1.3 RESEARCH GAP

In the absence of specific workload measurement techniques, this study was undertaken with the aim of developing a mental workload measurement and prediction tool that could be applied in the train control environment. The research problem is the objective measurement and quantification of TCOs' mental workload in order to compare the various train control centres in terms of workload and to predict the workload at a particular centre, given a certain set of conditions and variables.

The intended outcome of the study is to develop an objective method by which the mental workload of Train Control Officers (TCOs) can be measured and to develop a task-related index that can serve as an indication of actual or potential mental workload imposed on operators in specific work environments. In the railway environment, as in any safety-critical environment, mental overload can have potentially dire consequences.

The population, on which the mental workload assessment methodology was developed and tested, is Train Control Officers at Spoornet, the South African Railway Operator.

The objective of this research is the improvement of the performance and reliability of TCOs through the prevention of mental overload and ultimately therefore improving rail safety. In order to achieve this objective the following tasks were identified:

- Identify all possible factors that could contribute to the mental workload of TCOs.
- Develop methods to measure these factors.
- Verify the measurements to ensure validity and reliability.
- Develop an index to measure and predict mental workload that would stand the tests of objectivity and scientific scrutiny, and above all, contribute to a safe working environment.

1.4 JUSTIFICATION FOR THE STUDY

The proposed deployment of Radio Train Order (RTO) by Spoornet management as a means of train traffic control in specific areas has brought about a renewed focus on and examination of RTO by line management responsible for effecting changes in train control.

As a result, the shortcomings of the existing system have been highlighted. The main concern was that a previously set norm for the number of radio train orders that could be handled safely by TCOs existed. No documentation could be found that indicated the rationale for its development and the reasons for including only certain variables. The assumption was made, based on the elements in the formula used to calculate the norm, that it relates to the limitations of the technology rather than to those of the operators. There was also considerable uncertainty about the validity of the norm, as it was based on a purely mechanistic formula. Furthermore, there was concern regarding the replacement of an existing train control system with RTO while not knowing what the existing norm referred to or what it aimed to limit. It is a known fact, as explained before, that the decision to utilise a specific train control system is determined by a variety of factors. This automatically gives rise to the question: If the train control system is replaced by a different system (such as RTO), what guarantee is there that the new system will be suitable for the specific section? The underlying question from a human factors perspective is: Considering the change in workload, what guarantee is there that human operators in specific train control centres, with their own set of task demands, can safely operate RTO? In addition to this, what would be the safe norms within which to operate?

The following formula was used to calculate the workload of TCOs in the RTO environment. As far as can be ascertained, no studies were performed to prove the validity of this method.

Elements of the existing workload assessment method are:

i. Number of trains
ii. Number of train order stations
iii. Calculated over a 24-hour period

The formula used in terms of the above-mentioned method to calculate the workload of TCOs is:

$$\frac{\textbf{Number of Trains} \times \textbf{Number of Order Stations}}{\textbf{24}} \leq \textbf{100}$$

The norm that was set for a high workload was that if the calculated ratio exceeded 100, the workload was arbitrarily considered too high for that TCO. No rationale was provided for the threshold level of 100.

When analysing this formula, some issues were detected that support concerns about regarding the validity of the index:

- The fundamental problem with the current method is that it is calculated over a 24-hour period and does not consider the number or length of shifts. A 24-hour period can consist of two 12-hour or three 8-hour shifts.
- Different operators work different shifts and the number of trains are not evenly distributed over the shifts. The index above could measure some form of workload but it is not indicative of a single TCO's mental workload unless the number of trains and order stations are calculated per shift.
- As far as could be ascertained, the norm of 100 for a high mental workload has not been validated in any way. It is still not clear what constitutes a high workload.
- It is also not clear whether all the relevant factors that contribute to mental workload, such as a TCO's experience and the difficulty factors of the section they control have been included in the index.
- When applying this formula, a section with four trains over a 24-hour period and 100 order stations, and another section with 100 trains over a 24-hour period and four order stations gave the same workload rating. The perceived workload of the two scenarios was tested with a panel of expert users of RTO. They unanimously agreed that 100 trains with limited crossing places (four) compared with four trains and 100 possible crossing places would impose workloads on operators that are comparable to the extremes on a continuum.
- No scientific proof exists that existing method is valid (that it actually measures mental workload).

With analysis of the existing workload assessment method being used for train control officers at Spoornet, it becomes clear that its predictive, construct and face validity are questionable. The concern is raised that, while this method attempts to measure workload in the train control environment, it is possible that it does not measure the most pertinent factors that contribute to operator mental workload. Since the existing method is used as the norm for safe operations, this situation is potentially problematic for the employer, especially considering the legal requirements of employers in terms of employee and public safety (Occupational Health and Safety Act, 1993).

The Federal Railroad Administration (FRA) in the United States of America has performed a safety assessment of train dispatchers in 1987/88. During this exercise they collected data on the number of trains handled and the number of authorities issued by individual dispatchers over the course of a shift. This assessment considered the same elements as the formula used by Spoornet for calculating the workload at RTO centres. The FRA determined this to be an imprecise method of measuring dispatcher workload since it did not take into account the varied tasks that dispatchers must perform to move a train across the assigned territory (Popkin, Gertler and Reinach, 2001).

This study was conducted in an RTO control centre environment. Other train control systems in use on Spoornet infrastructure are the following:

- Colour Light Signalling System (Local Panel and Centralised Traffic Control [CTC])
- Radio Train Order (RTO) and Track Warrant System (TWS) (individually and in combination)
- Van Schoor System
- Semaphore Signalling System
- Telegraph Orders System
- Wooden Train Staff and Paper Ticket System

RTO-controlled sections interface with these other systems, which add to the complexity of the TCO’s task.

In the RTO environment, train orders are authorisations given by radio to train drivers to proceed to a predetermined point. Normally, these points are 14 kilometres (or less) apart. Contact is usually initiated by the train driver who needs to obtain an authorisation to pass a specific kilometre point and proceed to the next specified point. RTO is operated on a single track and all communication is conducted over the radio. All radio communications and telephone conversations are captured on a voice logger, which logs the date, time, content and duration of every conversation. The TCO’s function involves, among other things, planning daily train movements based on a given train schedule, making out authorisations, communicating with train drivers and dealing with telephone queries (normally of an operational nature). A track warrant is exactly the same as a radio train order, with the added safety feature of capturing train movements and authorisations on computer. The authorisation is then checked and approved by the computer. Authorisations which conflict are blocked by the computer and an authorisation to proceed will not be issued. TWS acts as a safety check for the actions of the TCO.

There are a variety of errors that can be made by TCOs. In the RTO environment these can be the misunderstanding of the position of a train, the issuing of an incorrect authorisation, or incorrect planning. These could lead to trains not crossing at the planned crossing places, and thus possible collisions.

With less equipment that could potentially be vandalised, RTO in combination with the added safety feature of the TWS seems to be favoured as the train control system for the future in South Africa. This would be especially true in areas where CTC is considered to be too vulnerable and expensive. It is foreseen that RTO will be expanded and introduced in more areas over the rail network in South Africa. For this reason an objective mental workload index is required in the train control environment. This type of index may prove valuable in other environments too, such as air traffic control and control room monitoring activities (i.e., nuclear installations).

In a paper delivered at the International Conference on Occupational Ergonomics in 1984, Moray (as cited in Meshkati, Hancock, Rahimi, and Dawes, 1995) made the following comments on the measurement of mental workload. These comments reflect the very same issues that Spoornet is faced with and that this study aims to address: (The italics are the author’s own emphasis.)

> Paradoxically, the fact that the operator seems to be doing less and less as processes become automated has resulted in a greater need for a measure of workload.
>
> If all operators showed graceful degradation under load, the problem would not arise. But while some operators allow their performance to degrade progressively

> as the load builds up, others appear to compensate for the effect of the load and show no change until a final catastrophic breakdown occurs. *In terms of safety requirements, if for no other reason, it would be highly desirable to have a metric for load which would allow one to detect the approach to the breakpoint independent of observable changes in performance. Such a metric does not exist***.**

This statement, by highlighting that there is still no workable index that measures mental workload in practice, clearly confirms the need for this study. The study areas of fatigue, stress and the other factors that contribute to mental overload have been studied and researched extensively but mental workload and specifically the measurement thereof, has many research gaps.

1.5 AIMS OF THE RESEARCH

The current unsatisfactory and potentially unsafe situation that Spoornet is faced with in determining the workload of TCOs working with RTO is not an isolated one. Similar scenarios are evident throughout the literature and the related comments echo the urgent need to develop a methodology with which to measure mental workload. This methodology needs to be validated and based on existing theories to ensure safe and accurate operations.

In short, the objective of the envisaged Mental Workload Index is to create a management tool that is proven to be valid and reliable and that can be applied to classify different lines or sections (this refers to the length of track that is controlled by a specific TCO before another TCO in the next train control centre takes over the movement of the train – the distance could range from under 100 kilometres to about 1000 kilometres) in terms of the mental workload these would impose on operators. This classification would assist in planning and allocating human and other resources, where and when necessary, and in taking the preventative steps necessary to minimise the risk of operator overload.

1.6 HYPOTHESIS AND RESEARCH QUESTIONS

The scientific hypothesis that will be tested by this research is:
The objective mental workload index as developed in this thesis can differentiate between high and low imposed workload train control centres.

In the process of developing the Mental Workload Index and eventually comparing the calculated workload at the various train control centres with physiological measurements associated with mental workload, the following research questions will be addressed:

- Which operator tasks could potentially be associated with mental workload?
- Which moderating factors, which would either increase or decrease mental workload should be considered and how can their respective moderating effects be determined?
- Which parameters should be considered to develop a measure of a mental workload that is completely objective and requires no estimation of experienced load by the operator?
- Can the proposed index differentiate between workload levels that could potentially be experienced at different train control centres?
- Which physiological measurements provide an objective assessment of mental workload?

1.7 OUTLINE OF THE THESIS

The contents of the thesis is organised in the following manner:

Chapter 1 – Introduction
In the Introduction the problem statement, research gap, justification for the study as well as the aims and the scientific hypothesis are provided.

Chapter 2 – Literature study
This chapter covers the elements related to mental workload research that could be found in the literature namely, the definition of mental workload, other related subject areas such as stress and fatigue, and the variables that affect mental workload. Mental workload

assessment techniques are discussed in detail and in closing, the study is contextualised in current railway related research.

Chapter 3 – Methodology

This chapter represents the essence of the study and provides a detailed account of the development process of the MWLI.

Chapter 4 – Results

The results of the verification study are discussed in this chapter. Key results are discussed but the full report is attached as Appendix D.

Chapter 5 – Discussion

The results are linked to the original hypothesis and possible applications of the MWLI are discussed.

Chapter 6 – Conclusions

Closing remarks are made and the limitations of the study as well as recommendations for further research are made.

Appendices

Appendix A - Cohen's Perceived Stress Questionnaire

The questionnaire was used in the pilot and verification studies to determine the perceived stress of Train Control Officers.

Appendix B - Timeline Analysis Template

The timeline analysis was used to capture the activities of Train Control Officers during the verification study.

Appendix C - Pilot Study Report

The study was undertaken to determine whether the physiological parameters rendered useful results.

Appendix D - Report on Validation Study

The verification study was undertaken to correlate the calculated Mental Workload Index with physiological parameters.

CHAPTER 2

LITERATURE STUDY

The elements of mental workload that are addressed in this chapter relates to the definition of the concept, its association with other related operator states and the variables that influence mental workload. An operational definition of mental workload is developed, based on the existing literature and in the context of this definition, the various mental workload assessment techniques are discussed. In closing, relevant railway related research is referred to in order to create the context for the study.

2.1 THE DEFINITION AND DESCRIPTION OF MENTAL WORKLOAD

2.1.1 The concept *mental workload*

The term *workload* covers a broad spectrum of human activity. *Mental workload* focuses on activities that are primarily mental or cognitive in nature and may involve physical coordination, but excludes activities resulting only in muscular fatigue.

The practical importance of the concept of *human mental workload* was established several decades ago during the investigation of human-machine systems such as ground transportation, air traffic control and process control. The theoretical development of the field can be traced back to a NATO conference and the subsequent text, Mental Workload (Moray, 1979). Since that seminal volume, many studies have been conducted on the theoretical underpinnings, assessment techniques and real-world implications of mental workload in a variety of work domains (Parasuraman and Hancock, 2001).

Mental workload should not be treated as a pure science – it is an applied science or technology. It should be viewed as a theoretical construct that is defined differently by engineers and psychologists. Engineers would typically put emphasis on operational definitions based on time available to perform a task, while psychologists would tend to emphasise the information processing aspects and define it in terms of measures related to channel capacity and residual attention. Physiologists may rather emphasise considerations of operator stress and arousal (Wierwille, 1988).

The intention of this study is to integrate these views into an operational definition in order to assess the relevance and effectiveness of techniques that objectively measure mental workload. From a review of the literature it is clear that the multifaceted nature of human mental workload has been emphasised by various researchers. The literature presents the reader with a maze through which careful navigation is necessary in order to attain clarity and avoid further confusion. In the search for a definition of workload, it is clear that different perspectives have led researchers to define workload in various ways. A number of definitions are reviewed here in order to develop a working definition for the purposes of this study.

Not only is there no single, commonly accepted definition of workload but there are many conflicting concepts of what workload comprises, and often the term is used without any definition at all. There is also the difficulty of differentiating workload from stress, which has many features in common with workload. Often the terms workload and stress are used interchangeably (Meister, 1985).

When the concepts of fatigue and situational awareness are added to those of workload and stress, one realises that the territory is complex and problematic, especially for newcomers to the research field. Researchers are rarely confronted with only one of these issues. All of these issues present themselves in the environments in which operational research is performed. These environments are always safety-critical in nature. One only has to think of the environments where mental workload research originated, namely air traffic control, aircraft handling and aerospace operations to realise that these operators are all exposed to stress and fatigue (most of these are 24-hour operations), they have to be situationally aware and they could potentially become mentally overloaded.

This section therefore aims to derive an operational definition of mental workload that is grounded in the research of the past three decades. The overview of the research on mental workload does not claim to be exhaustive but covers the most important conceptual work in the field.

In an attempt to create some order among the multitude of mental workload definitions, they have been grouped together based on the premise from which the definition was developed. As many of the definitions were based on more than one premise, this proved fairly challenging.

2.1.2 Mental workload as a function of capacity

Mental workload is often conceptualised as the interaction between task demands and the *capacity* of the operator. Thus the ratio between demand and capacity determines the level of workload. *Capacity* is determined by the operator's skills and training but may also be influenced by stressors such as fatigue, noise etc. Task demands are determined by the number of tasks to be performed, the amount of attention needed and the time available.

Human operators could be viewed as communication channels with limited and fixed *capacity*. It has been theorised that as long as total capacity limits are not exceeded, capacity can be divided at any time among concurrent processes and activities. Another view is that the capacity needed for a particular activity is expressed as the time taken by that activity, with a more complex activity taking more time. In this case, mental load then equals input load, which can be largely defined in terms of task variables and can be directly measured according to the time taken to perform activities. A considerable amount of capacity may remain unused in a single reaction to task demands because only a limited number of mechanisms may be involved in that reaction. It can be argued that the more complex the task, the more mechanisms are involved and consequently the more capacity is used (Senders, 1979).

Capacity does not remain constant however. Available capacity increases with the task demands as a result of physiological mechanisms related to arousal. This reflects the physiological components and it can be argued that mental load equals the effort needed to transform the actual task demand into an adequate work result. Factors such as 'willingness to spend capacity' may moderate the amount of effort operators are prepared to spend on a specific task (Moray, 1979).

Workload measurement is the specification of the amount of capacity used. In this definition workload is not solely task-centred. Mental workload depends on the demands in relation to the amount of resources operators are willing or able to allocate and is therefore a relative concept (O'Donnell and Eggemeier, 1986).

Related to the concept of capacity but referring to operators as adaptive agents, Parasuraman and Hancock (2001) propose that mediating factors should be considered rather than assuming that mental workload is an intervening construct that reflects the relation between the environmental demands imposed on operators and the *capabilities* (or capacity) of operators to meet those demands. Workload may be driven by the taskload imposed on human operators from external environmental sources but not deterministically so. This is because workload is also mediated by individual responses to the workload and varying skills levels, task management strategies and other personal characteristics. Although such definitions immediately suggest the notion of adaptation by human operators, most studies have failed to explicitly examine this aspect of workload. The majority of studies have been concerned with empirical evaluations of the effects of various taskload factors on various measures of mental workload – whether based on performance outcome, physiological response or subjective report.

2.1.3 Mental workload defined in terms of *experienced load*

A simplistic definition of workload is that it is a demand placed on humans. This definition attributes workload exclusively to an external source. However, workload can be better defined in terms of experienced load. With *experienced load*, workload is not only task-specific but also person-specific. Besides individual capabilities, the motivation to perform a task, the strategies applied in task performance, mood and operator state also affect experienced load. Therefore, workload depends on the individual, and as a result of the interaction between operator and task structure, the same task demands do not result in an equal level of workload for all individuals (De Waard, 1996).

Workload has been described as multidimensional because it can be viewed either as an input or as an output or consequence. As an input, workload is represented by stimuli that load the operators in the sense that they are caused to bear a burden. As an output, workload affects not only performance but, when operators are part of a larger system, it impacts (usually negatively) on the system itself. As an intervening variable, workload is the operators' internal experience of difficulty and discomfort, their recognition that they are experiencing a load and the strategies they employ to overcome this. It can be said that as an intervening variable, workload alters the human information-processing system by allowing the operator to adopt strategies to cope with the experienced workload. Workload can thus be viewed as a feature of the system (even when it is the operator's own incapacity) that forces operators to work harder. It can also be viewed as their feeling of being stressed and of having to work harder. The effect of these factors may cause them to make errors (Meister, 1985).

Mental workload also represents the cost incurred by a human operator to achieve a particular level of performance. An operator's subjective *experience of workload* is the product of many factors, in addition to the objective demands imposed by the task. Thus, workload is not an inherent property but rather the product that emerges as a result of the interaction between the requirements of a task, the circumstances under which it is performed, and the skills, behaviours and perceptions of the operators themselves (Hart and Staveland, 1988).

This internal experience of difficulty and discomfort is the basis of the stress created by mental workload. This stress is in most cases, especially for TCOs, a time-related factor.

2.1.4 Mental workload defined in terms of *time load*

The definition below relates mental workload to time but then considers, as an additional factor, limited capacity which results in competition for attention:

i) Tasks must be performed within a certain length of *time*. The degree of workload is the percentage of that time which the operator actually has to perform those tasks. This concept focuses on time availability. A high degree of workload is experienced when the task requires most or all of the available time.

ii) The operator has limited capacity, usually conceptualised in terms of attention. When

> the operator must perform multiple tasks within the same *time period*, there is competition for the operator's attention. That competition 'loads' the operator (Meister, 1971).

In a similar vein, Reid and Nygren (1988) propose that mental workload should be viewed as a multidimensional construct that can be explained largely by three component factors: *time load, mental effort load* (which relates to capacity) and *psychological stress load* (which seems to be more related to the consequences of the previous two). Again, as with Meister's definition, workload is defined in terms of two factors: time and capacity.

Time load involves both the time available for a task and task overlap. If the time required to perform a task exceeds the time available, the operator encounters a time-load problem. If the tasks that the operator has to complete start to compete for time resources, the operator will be forced to evaluate the tasks for priority and allow the performance of some tasks to deteriorate and/or their completion to be delayed.

The channel capacity concept deals with task factors such as difficulty, complexity and required effort. This dimension assumes that the human operator has limited capacity. Performance of one task may consume a certain amount of an operator's resources, while another task may consume other resources. The implication is that resources that are not expended in task performance are either held in reserve to be used for other tasks or used for extra effort in accomplishing the current task. This dimension therefore refers to *mental effort load*.

The third component factor deals with the concept of *psychological stress* and encompasses a number of operator variables such as motivation, training, fatigue, health and emotional state. This dimension is broadly defined as anything that contributes to an operator's confusion, frustration or anxiety and is referred to as psychological stress load.

2.1.5 Mental workload as a function of *demand and supply*

Workload can be defined in terms of the relationship between *resource supply and task demand*. It is argued that operator workload is directly related to the extent to which the tasks performed by the operator utilises the limited resources (Wickens, 1992).

Figure 2.1, reproduced from Wickens and Hollands (2000), shows the supply-demand relationship between the important variables in the workload model. The resources demanded by a task are shown on the horizontal axis. The resources supplied are shown on the vertical axis, along with the level of performance. If adequate performance of a task demands more resources from the operator than are available, performance will break down, as shown on the right of the figure. If, however, the available supply exceeds demand, as shown on the left, then the excess expresses the amount of reserve capacity. According to this model, changes in workload may result either from fluctuations of operator capacity or from changes in task resource demands.

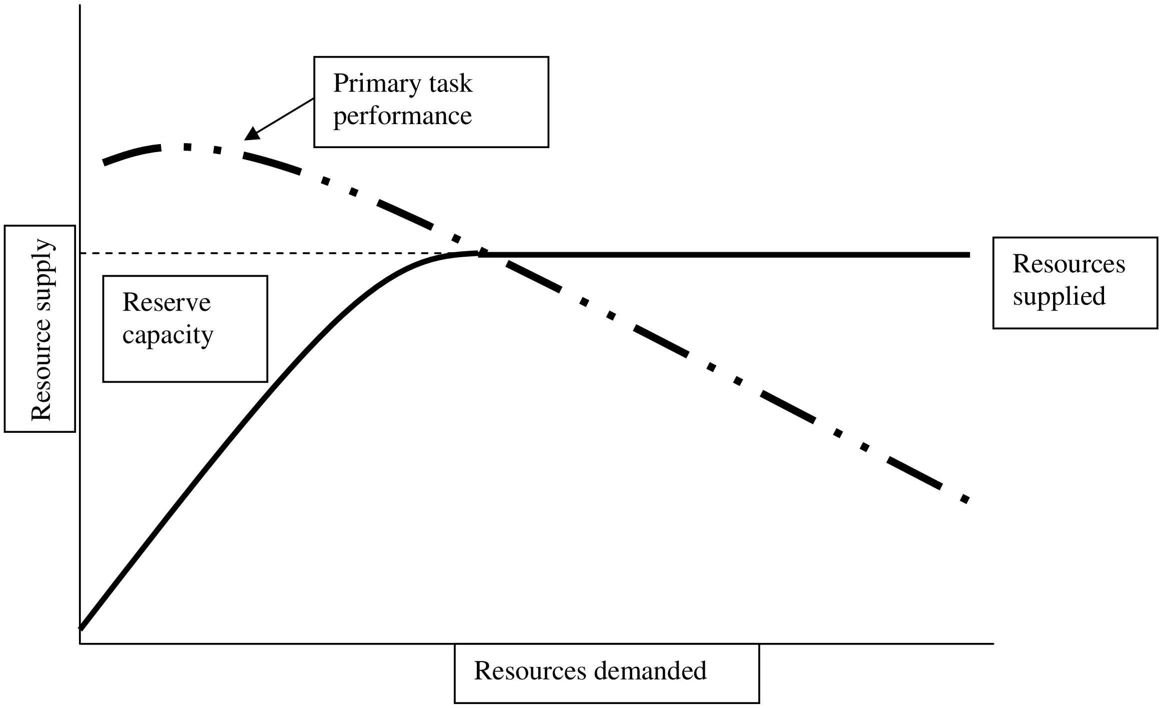


Figure 2.1: Schematic relationship among primary-task resource demand, resources supplied and performance (Wickens and Hollands, 2000).

2.1.6 Mental workload versus taskload: an argument for a continuum

The term *taskload* is often encountered in the literature on mental workload. In some instances the term is used to refer to the task demands placed on the operator but it is often used as a synonym for mental workload. It would seem more likely that the distinction between the two is a continuum rather than an absolute distinction.

Simply stated taskload is what the operator is confronted with in the execution of duties i.e., the number of tasks that have to be carried out. Mental workload also takes into consideration what the operator does to cope with the work. Workload is often defined as the interaction between the demands of the task and the ability of the operator to meet those demands (Reinach, 2001).

Hilburn and Jorna (2001) distinguish between the two concepts as follows: Taskload is the demand imposed by the task and mental workload is the operator's subjective experience of that demand.

A number of studies have attempted to identify taskload indexes for Air Traffic Controllers (ATC). Out of many prospective taskload indexes, the number of aircraft under control (i.e., traffic load) has shown the clearest predictive relationship to workload measures. However, traffic load alone does not accurately capture the total load imposed by what is happening in the airspace. Factors such as skill, training, experience, fatigue and other stressors all mediate the relationship between task demands and workload experienced by the controller.

The distinction between taskload and workload in the ATC context can be depicted graphically as follows:

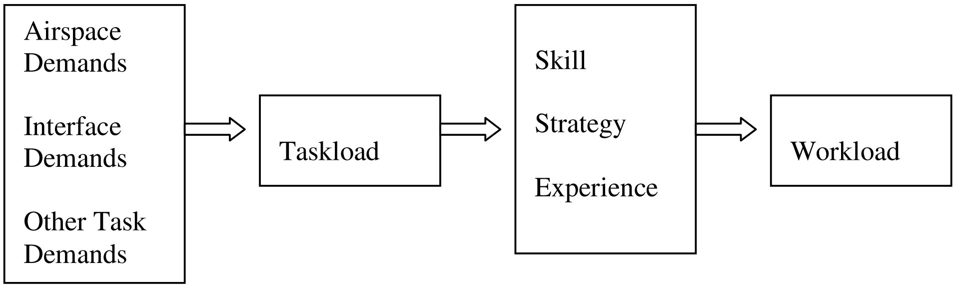


Figure 2.2: Taskload and workload determinants in ATC (Hilburn and Jorna, 2001).

The ratio between demands and capacity therefore determines the level of workload. Capacity is determined by the skills and training of the operator but may also be influenced by stressors such as fatigue, noise etc. Task demands are determined by the number of tasks to be performed, the amount of attention needed and the time available (Hancock and Meshkati, 1988).

The above distinction between mental workload and taskload is very similar to that of Hilburn and Jorna (2001) and shows that something interacts with the task demands (capacity) in order for taskload to become mental workload.

To place taskload and mental workload along a continuum may seem mechanistic but this means of definition would appear to add both clarity and understanding to the construct of mental workload.

2.2 MENTAL WORKLOAD AS RELATED TO OTHER OPERATOR STATES

Although the states of workload, stress and fatigue may occur simultaneously and are sometimes difficult to distinguish from one another, they should not be regarded as synonymous and their differences should be clearly understood. A proper distinction between these concepts is not only important for theory building but also for the restructuring of the work environment (Gaillard, 2001).

The confusion between these concepts originates in their poor definitions. As discussed in the previous paragraphs, mental workload may refer either to the objective demands imposed by the task (e.g., complexity or pacing) or to the subjective judgement of the operator with regard to the task demands. In most theories, workload refers to the limitations in the information processing capacity of the operator. It may, however, also encompass feelings of work pressure, which have a more emotional connotation.

2.2.1 Workload as related to stress

The concept of stress has a large variety of meanings:

- An *input* variable referring to either work demands (difficulty and time pressure), emotional threat (accidents) or adverse environments (noise and sleep loss).
- An *output* variable referring to a pattern of behavioural, subjective and physiological responses, often labelled *strain*.
- A *state* in which we feel strained, pressurised and threatened on the basis of a subjective evaluation of the situation.
- A *process* that gradually results in a dysfunctional state, degrading the work capacity and the potential to recover from work.
  (Gaillard, 2001).

Load is defined as the demand placed on the operating resources of a system. Stress and strain are indications of the effects of workload on individuals. Mental workload does change an individual's mental shape in terms of the resources they make available. Physiological variables change with mental activity and therefore seem to be well suited as indicators of changes in mental workload demand (Rasmussen, 1979).

Workload should not be confused with stress. Some of the negative effects of stress on performance include narrowing of the attention span, forgetting the sequence of actions, incorrectly evaluating situations, slow decision making and failing to carry out decisions (Meshkati and Loewenthal, 1988).

Hart and Bortolussi (1984) draw the following distinction between workload and stress in their assessment of sources of workload. Workload was evaluated as the change in effort required when dealing with tasks, whereas stress was regarded as an experience that is the consequence of changes in the tasks. However, the results of their study indicated closely correlated estimations of workload and stress.

2.2.2 Workload as related to fatigue

Fatigue is interpreted as a response of mind and body to a reduction in resources resulting from the execution of a mental task. It is also considered to be a warning of the increasing risk of performance failure. Fatigue may also refer to a subjective complaint that encompasses a general feeling of lack of energy that is not necessarily related to the amount of work (Gaillard, 2001).

Although two types of fatigue are referred to, only passive fatigue is relevant in the train control context:
Active fatigue results from continuous and prolonged, task-related, perceptual-motor adjustment.
Passive fatigue relates to system monitoring with either rare or no overt perceptual-motor response requirements. Closely related to vigilance, this form of fatigue develops over a number of hours during which individuals appear to be nothing at all (Desmond and Hancock, 2001).

Fatigue represents one form of stress. Fatigue is considered as the product of some aspect of environmental input on the individual, such as heat, noise, poor ergonomic design of the work environment and number of hours worked. Fatigue is evident in performance decrements of operators.

This is one of the reasons why *shift worked* has been included as a mediating factor in the MWLI. This information can be collected objectively and it represents an input stressor on the individual.

2.2.3 Workload as related to situational awareness

Situational awareness has not received as much attention as stress and fatigue, but it is worth mentioning in the mental workload context.

Situational awareness is defined as the perception of elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future. It is a uniquely cognitive phenomenon which supports action but is not part of the action itself. It applies to a dynamic, changing environment, and the nature of the situation of which awareness is maintained, needs to be specified.

There is an interaction between workload and situational awareness. Maintaining a high and accurate level of situational awareness is a resource-intensive, cognitive process. Thus, on one hand, one cannot gain accurate situational awareness without expending resources, and this may in turn compete with other concurrent cognitive tasks. On the other hand, heavy concurrent task demands may divert resources from the maintenance of situational awareness.

This relationship between situational awareness and mental workload is mediated, to a certain extent, by the degree of operator expertise or skill. Skilled operators can generally preserve situational awareness with lower resource cost (Wickens, 2001).

The terms mental load, fatigue and arousal are often viewed as synonymous with stress. Despite attempts to give unique definitions to these terms, it is difficult to see what their distinguishing features are.

The interrelatedness of the concepts of workload, stress and fatigue is clearly complex, often confusing and a potential pitfall for researchers. From the outset of this study it was decided to simplify the approach to workload as far as practicably possible but not to lose sight of the other related factors. By adopting this approach, this study could be criticised for simplifying the concept of workload, a calculated risk that had to be taken in order to provide a user-friendly, fit-for-purpose solution to an operational challenge.

2.3 AN OPERATIONAL DEFINITION OF MENTAL WORKLOAD

2.3.1 Comments

Considering the wealth of information about mental workload that has been generated over the past three decades and the variety of definitions from different perspectives, it is disappointing that consensus has not been reached among researchers as to how mental workload should be defined. This makes it difficult for new-generation researchers. Moray's 1979 foundation, as the starting point, has remained relatively unelaborated upon. Although this is an excellent foundation, there is an expectation that advances in terms of reaching an agreement on the definition of mental workload would have been made.

As far back as 1988, Wierwille (1988) expressed the opinion that mental workload research, having made initial strides, appeared on the verge of becoming mired in its own details. It is important that momentum in this research is not lost. It is doubtful whether, in terms of finding a consensual definition of mental workload, there has been any progress since this statement was made.

As a final comment on the definition of mental workload, it is important to continue questioning the very concept of mental workload. After many years and many hundreds of empirical investigations, there is still no satisfactory, consensual definition of workload. A deeper understanding of mental workload is essential for the following reasons:

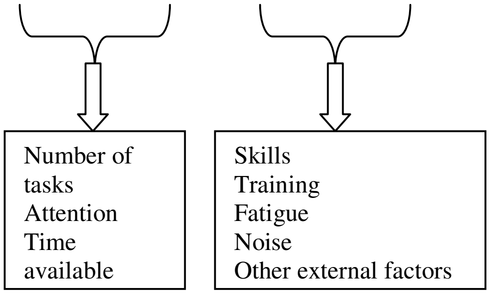
- Theoretical motivations – we seek to understand and relate the phenomenon of mental workload to related cognitive constructs such as attention, effort and resources.
- Practical motivations – the nature of work in the developed world has changed from being physical to primarily cognitive. The 21st century is an era of intelligent automation and there is a need to assess the load of such information-based work. It is important to know how much mental load is too much and may prove hazardous or stressful. There is also a need to know how little is too little, so that individuals are sufficiently challenged to sustain useful levels of output.
- Technological motivations – mental workload must be assessed as an indicator for use in complex technical systems, to dynamically change the nature or the demands of the work at hand. As almost all forms of work now involve human-computer interaction,

this strategy becomes progressively more feasible, and indeed crucial, for complex systems.
(Parasuraman and Hancock, 2001).

2.3.2 Synthesis of definitions

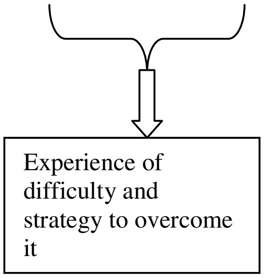
When the four categories of definitions are considered, the following key elements are evident:

- Mental Workload is the interaction between: *task demands* and *capacity*



- *Experienced load* – is an intervening variable

Input ⇒ Intervening variable ⇒ Output/Performance



- *Time load* – relates to capacity and stress

- *Demand and supply* - relate to task demand and resource supply

These elements fit neatly into the very comprehensive and cohesive model proposed by Meshkati (1988) (see Figure 2.5). This model maps the major interrelationships of the interacting variables of mental workload. Meshkati's model is the most comprehensive single model of workload found in the reviewed literature and it ties together most definitions of mental workload.

Meshkati's model consists of two sections – Causal Factors and Effect Factors – each with two primary component groups.

- The two primary groups of Causal Factors are the following:
    - task and environmental variables and
    - operators' characteristics and moderating variables.
- The two primary groups of Effect Factors are the following:
    - difficulty, response and performance variables and
    - mental workload measures.

This model has been found to be very helpful in integrating the various concepts relating to mental workload but is somewhat complex to use when it is necessary to explain the concept of mental workload to an uninformed audience. The model proposed by Jahns (as cited in Johannsen, 1979) appears to be a good foundation and one that is both sufficiently comprehensive (with a few additions) and easy to use.

Jahns (as cited in Johannsen, 1979) has proposed a concept for the assessment of workload that divides the broad area of human operator workload into three functionally relatable attributes:

Input load ⇨ Operator Effort ⇨ Performance

In terms of this model the following can be said:

- Major sources of input load may be separated into three categories: environmental, design-induced or situational and procedural. Environmental variables are noise, vibration, temperature, etc. Design variables are, for example, the characteristics of displays and control devices, crew station layouts and vehicle dynamics. Procedural variables include briefing and instructions, task sequencing and task duration.
- Operator effort depends on a number of factors including the input load and the performance requirements of a given task, as well as internal goals and motivation. The operator state depends on many factors. Relatively stable factors would be

psychophysical characteristics, general background and personality. Fluctuating factors would be experience, motivation and attentiveness.

- Internal performance criteria are maintained by human operators and influence their tolerated error levels. These criteria depend on factors of the operator state (e.g., motivation), performance requirements and instructions.

Workload is the umbrella concept which includes input load and operator effort.

Using the elements in Jahns' model as the foundation and relating it to the definitions of mental workload referred to earlier in this chapter, the following working definition of workload is proposed as a basis for the development of the MWLI.

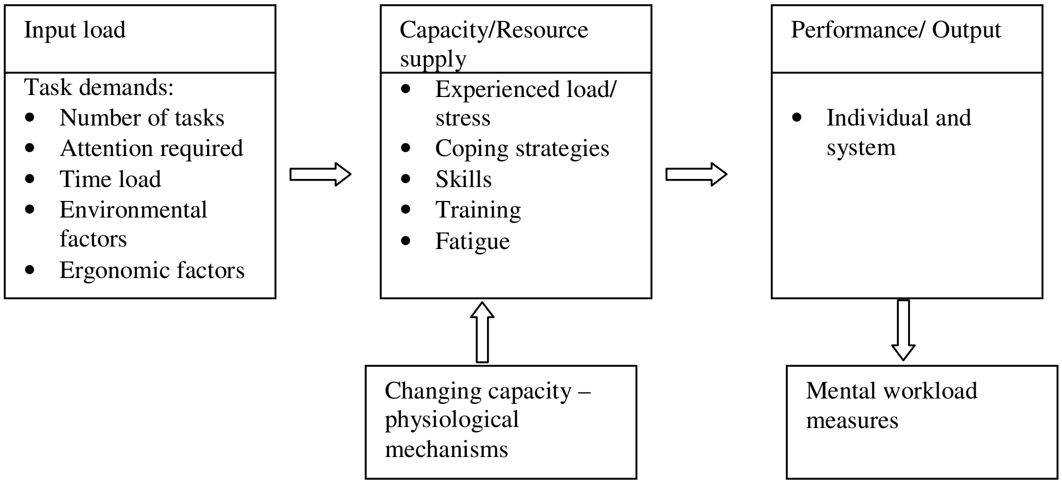


Figure 2.3: Flow diagram of an operational definition of mental workload

2.4 VARIABLES THAT INFLUENCE MENTAL WORKLOAD

Mental workload, like fatigue and occupational stress, is one of the factors in the human-machine/system interface that needs to be managed in order to sustain safety and productivity. Now that an operational definition for mental workload has been established, it is necessary to briefly refer to those variables that could influence mental workload.

There are two categories of variables that will be discussed:

- individual factors
- task factors

The Ergosystem provides a framework for relating the factors that influence mental workload. For any system to function effectively, all the elements must be integrated in a functional manner. The nature of any system is such that any change in a sub-system will result in disequilibrium of the system.

Human
- Biological
- Psychological (Capacity/Motivation)
- Skills/Training

System/Machine/Equipment
- Physical Stressors
- Psychological Stressors (Fatigue)

Human-Machine/System Interface
- Task
- Design

Environment
- Occupational Hygiene Factors (Noise)
- Psychological Factors
- Organisational Factors/Culture

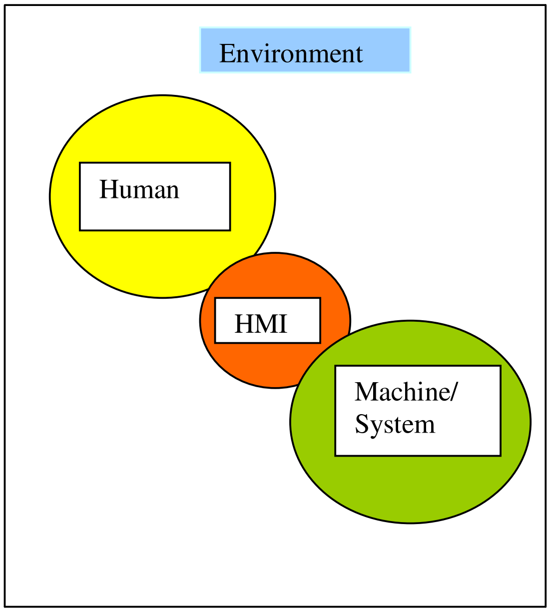


Figure 2.4: The Ergosystem (Adapted from Van Tonder, 1999)

2.4.1 Individual factors affecting workload

A common feature mentioned in the assessment of mental workload is the differential sensitivity of measurement methods to the factor individual differences or personality traits. This has been acknowledged as an influential variable. It manifests itself in the subjective assessments and physiological responses of operators (Meshkati, 1988).

In this regard, it is worth noting the comments of a few experts. According to Firth (1973), in the real-life working environment individual differences in operators' characteristics greatly influence an individual's information-processing capacity. These differences arise from a combination of past experience, skill, emotional state, motivation and the estimation of risk and cost that may be involved in performing a task.

Meshkati and Loewenthal (1988), studied the role of personality in performance decrement and attention. Their results indicated that individuals who scored high on a distractibility scale (i.e., extroverts) found it difficult to maintain a uniform mode of performance. This group of subjects exhibited increasing lapses in attention, while the introverted group failed to show any evidence of a decline in attention.

Damos (1988) has done a comprehensive study of the research available on individual differences in the subjective estimates of workload. She has identified the following variables as significant:

(a) Personality traits and behavioural patterns

Little research has been conducted that examines the relationship between mental workload and personality traits.

To date, the only behavioural pattern investigated is the Type A coronary-prone behaviour (Friedman and Rosenman, 1974). Individuals who have high scores on tests measuring coronary-prone (Type A) behaviour are characterised by an extreme sense of urgency. As a result, they prefer a more rapid work pace and tend to perform better on tasks that do not have a deadline than individuals who have low scores on tests of coronary-prone behaviour (Type B individuals).

It appears that, under dual-task conditions, Type A subjects generally experience less effort (or mental workload) than Type B subjects and the same or more effort under single-task conditions.

Meshkati and Loewenthal (1988) emphasised that cognitive style, i.e., learned thinking habits, may have an influence on behaviour, subjectively experienced workload and perceived task difficulty.

(b) Motivation

The effects of motivation on the differences between subjective workload assessments and performance have been reviewed by Vidulich (1988). It was predicted that although increased motivation would lead to better performance, it would do so by encouraging subjects to expend greater effort, thereby increasing the experienced workload.

(c) Mental Capacity

Mulder (1979) reviews the difference between the mental capacity expended by the main task and the total available capacity. It is assumed that the difference between the capacity of the operator and the load imposed by a specific task is called reserve capacity. The total available capacity is not constant but varies with the operator’s state of arousal. The more effort (i.e., processing capacity) an operator expends on the task, the less capacity remains available for other tasks or circumstances that may demand attention.

(d) Gender and time of day

Hancock (1988) reviewed the effect of two operator-specific factors, namely gender and time of day, on subjective estimates of mental workload. Significant differences were found between males and females. The study indicated a higher intolerance among female subjects for the repetitive and boring task examined. Should this represent a gender difference in tolerance to work under conditions of underload, this study may present an important finding, specifically from a selection perspective in safety-sensitive operations, such as train control or monitoring. No significant effects were found for time of day.

(e) Other Factors

Additional factors emphasised in the study of Meshkati (1988) include the importance of the individual's state of arousal, sensory capabilities and the level of training and experience. All of these factors may have a considerable influence on perceived mental load, which is reflected in workload measures.

In research it is imperative to state in the design phase how individual differences in mental workload assessment will be dealt with, rather than ignoring their potential effects. In this particular study a conscious decision to eliminate personal factors from the MWLI was made. This was done in order to obtain a predictive value per train control centre based on the 'average' operator.

2.4.2 Task variables that affect workload

An interesting concept that could only be found in one literature source is the notion that workload researchers have paid little attention to the sources of variance in the independent variables and, instead, have focused their attention and experimental conclusions on the dependent variables, i.e., measured indications of workload. The independent variables referred to are those task variables that affect mental workload (Cilliers, 1992).

The variables quoted here refer to the military aviation sector, and those that are potentially relevant to the train control environment will be discussed briefly.

(a) Task criticality

Task criticality in the military context can be viewed as a function of the task content and the effect of execution or non-execution on the mission or task as a whole. The criticality of a task increases with the unpredictability and the severity of the effect.

Analogous to the military aviation scenario, the various tasks or phases in the train control environment that constitute task criticality are the following:

- Planning of shift activities – by studying the daily train plan, TCOs can get an overview of which trains are in their section, which trains have to depart, at what time and from where and which other activities (maintenance and shunting) are planned for the day. These trains are then plotted on train schedule sheets and mapped according to the predetermined running times between sections. Crossing places are predetermined and mapped on the schedule.
- Communicating by radio with train drivers to obtain their positions and provide authorisations to proceed up to a predetermined kilometre point.
- Communicating with and monitoring conversations of track maintenance teams.
- Answering telephone calls from other staff (maintenance and commercial) regarding train movements.
- Writing of reports on activities/incidents such as train failures and shunting activities.
- Capturing of incidents such as derailments, infrastructure problems in RIMAS (Risk Information Management System).
- Termination of trains in their section (when a train enters the next train controller's section).
- Handing over tasks on completion of the shift.

(b) Environmental factors

According to Cilliers (1992), environmental factors also influence mental workload.

In the train control environment the following environmental factors are relevant:

- Design and layout of train control centres – seat, desk, ambient noise and temperature. Due to the multitude of possible variations of this factor and to the task team not considering this to be an important contributing factor, this was not included in the MWLI.
- Topography – the nature of the terrain and landscape through which the section runs (see description in Chapter 6, Table 6.4). This was included in the MWLI as variations in this factor impact on the ability of train drivers to keep trains running on time which could impact on the TCO's scheduling and planning.

(c) Amount of information and complexity

The amount of information to be processed by the pilot can be defined as a direct result of the tasks imposed by the mission as well as those tasks imposed by emergency conditions and the tactical environment. A portion of the information to be processed by the pilot is predictable and another portion is highly unpredictable.

In the train control environment a similar situation exists. The tasks of TCOs are highly structured and predictable when normal operations prevail. It is when an abnormal situation arises (i.e., an accident, derailment or communication failure), which requires quick problem solving and the use of discretion and prompt response, that the information to be processed becomes excessive and complex, which could lead to mental overload and increase the risk of errors.

(d) Task structure (High versus Low)

According to Cilliers (1992), task structure can be described as a function of design of the man-machine interface that influences the normal and emergency procedures, and the tactical environment which elicits the task response.

In both the aviation and train control environments, the task structure is high. This means that work is done in a proceduralised way. For instance, the issuing of an authorisation occurs according to a set step-by-step procedure. Train planning (which train has priority when crossing another train, for example) occurs according to a set of rules. It is only when equipment such as radio communication fails and abnormal working procedures are activated, that tasks become less structured, time pressure increases and the knowledge and discretion of TCOs are stretched. This is often due to the fact that TCOs have not been properly trained and/or the training has not been regularly refreshed when it comes to fall back or emergency procedures. When certainty and predictability decrease, workload increases.

(e) Time response

In terms of time availability for task execution, a distinction should be made between routine and emergency conditions. Under emergency conditions, time available will be a function of the number of emergencies (or out-of-the-ordinary tasks) to which TCOs must attend, as well as a function of the criticality of the emergency condition.

(f) Equipment and design

The radio train control environment is significantly less sophisticated than that of a cockpit but the effect of the environment on the workload of TCOs is no less important. The full spectrum of ergonomic factors (i.e., the adjustability of desks and chairs, ambient noise, temperature, illumination and the reliability of equipment) was considered when analysing the tasks of TCOs. Initially this factor was included in the MWLI but was later discarded by the task team due to the variability of train control centres and the potential for subjective interpretation.

(g) Task Novelty

It has been speculated that a reduction in task novelty will result in a reduction in the degree of workload experienced (Cilliers, 1992). This factor partially explains why more incidents occur when TCOs have to use fall back procedures, as may be necessary when equipment fails and the normal procedures cannot be followed. This provides a good rationale for regular refresher training for TCOs in abnormal working or fall back procedures, which will reduce the novelty of the task. The importance of this factor is emphasised by the inclusion of *number of years experience* as a moderating factor in the MWLI.

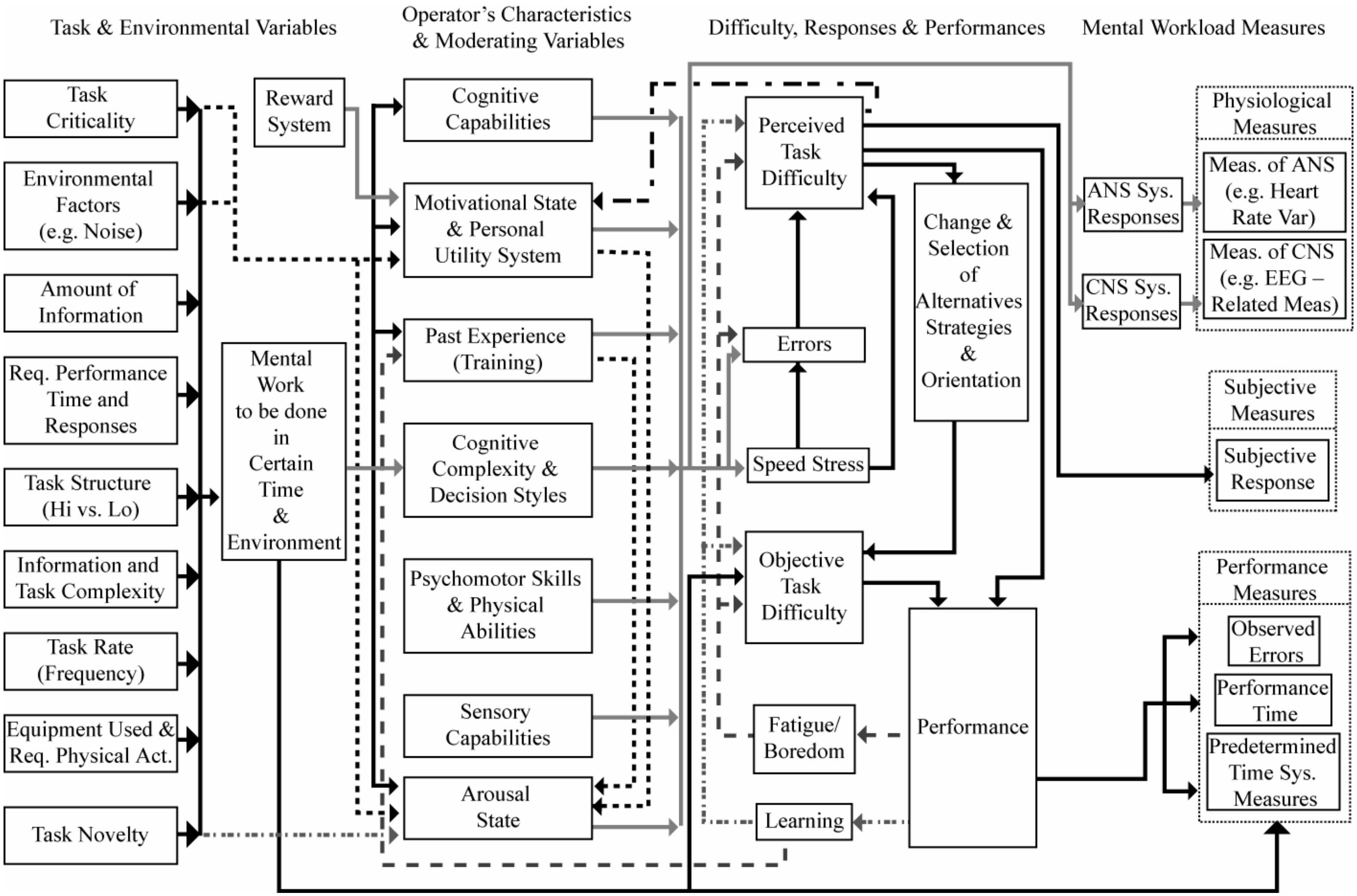


Figure 2.5: Major Components of Comprehensive Mental Workload Model and Related Assessment Variables (Meshkati, 1988)

2.5 APPLICATION OF WORKLOAD THEORY

As the previous sections indicate, the concept of mental workload is considered to be highly complex and cannot be simplified or reduced to a single dimension. This has lead to a situation where there appears to be an inability among the researchers in the field of workload to reach agreement on workload definitions, research procedures and the degree to which this information can be generalised from one context to another.

It has become clear that few researchers provide clear explanations as to the areas of application of their workload research. A contributing factor may be the fact that most of the studies were conducted for US and other military institutions, with military security classification, which has denied South African researchers access to this information. It appears however that much of the published research focuses on laboratory experimentation or simulated situations, and that the procedures used by researchers cannot be made applicable in a similar fashion in applied environments.

The study of Cilliers (1992) is the only South African study on mental workload that could be found. A search of the literature for the application of mental workload research principles or measurements rendered no returns. Railway companies that commission research, normally classify the information and it is either not available for public access or access is limited to a few published articles. Apart from a number of comprehensive Federal Railroad Administration (FRA) research reports, information on how other railway companies address the problem of mental overload in the train control or other operational environments is therefore limited to personal communications with individuals involved in this work. In addition, access to research is often dependent on membership of professional research bodies.

2.5.1 Application in safety operations

It would appear that individuals are most reliable under moderate levels of workload that do not change suddenly and unpredictably. Extreme levels of workload increase the likelihood of human error as a result of the operator’s inability to cope.

When workload is too low, the operator may become bored and lose concentration. This may also lead to increased errors.

2.5.2 Application in automation

Tsang and Johnson (1987) maintain that there are a number of reasons for incorporating higher degrees of automation in complex aircraft systems (or any monitoring and control environment). The most important reasons being the increase in the level of safety and the reduction of pilot (or operator) workload. Another important consideration is that it would free the pilot involved in the system from repetitive processes which could best be performed by automated devices. This would allow the pilot to concentrate on activities such as planning and decision making.

The closest that the Spoornet train control environment has come to the employment of automated devices is the use of computer-based systems such as TWS, which is programmed to disallow a planned train movement that is in conflict with a previously planned movement.

The following is concluded from the existing body of knowledge of mental workload:

- A variety of different definitions of workload are being used by researchers.
- Different perspectives (e.g., engineering, psychology or physiology) may lead to different definitions of workload. The importance of the context of workload measurement has been emphasised.
- Changes in the definition of workload have occurred over time. Initial emphasis was on the time dimension. As a result of increased research by psychologists, the information-processing aspects received more attention, and issues related to channel capacity and residual attention were highlighted. Psychologists have also investigated the concepts of arousal and stress as these relate to workload.
- Any attempt to decide which of the workload definitions is most accurate would appear to be unproductive. There is no empirical technique for providing a single definition of a multidimensional construct such as workload. Each definition is useful and contributes to the understanding of workload.

2.6 AN OVERVIEW OF MENTAL WORKLOAD MEASUREMENT TECHNIQUES

In this section a general overview of the existing mental workload measurement techniques is given. The criteria for the selection of mental workload assessment techniques are also discussed and a set of criteria that is applicable for this study is provided.

In the literature, the measurement techniques are divided into the following groups:

- measures of primary task performance
- secondary measures of spare mental capacity
- subjective rating techniques
- physiological indicators of mental workload

These same groups are used in the overview. All the techniques that could be found in sources available in the public domain will be referred to, rather than singling out a select few. When this study was initiated this was considered to be the most effective approach. Before the decision was made to develop the MWLI, the existing techniques were critically evaluated in order to find an existing tool that would address the requirements of this particular study. Each of the techniques is commented on in terms of its applicability and appropriateness for this study.

From the literature on mental workload assessment techniques, it is evident that a great deal of information has accumulated in the period between the mid 1970s and late 1980s. During this period, significant changes were occurring in the theoretical conceptualisation of workload itself. It became increasingly clear that workload was a multidimensional rather than a unitary construct. Workload became conceptualised as that portion of the operator's effort which was actually required to complete a task. This view clarified a very important fact regarding any attempt to measure the elusive construct workload. No single measure would ever suffice as the holy grail of workload measurements. The multidimensional nature of workload demanded that multiple measures be used to deal with the construct adequately. Thus while a single metric may assess a specific causal factor, it would be a mistake to demand, as a criterion of acceptance, that any such measure be used to generalise over all kinds of workload or even different levels of the same kind of workload (Wilson and O'Donnell, 1988).

During this period groundbreaking research was performed, which resulted in the development of some of the most widely used workload measurement tools, such as the Subjective Workload Assessment Technique (SWAT) and the National Aeronautical and Space Administration Task Load Index (NASA-TLX). Today these measurement tools are considered to be best practice or gold standards for subjective mental workload measurement. It is also evident from the literature that most of the development work of this period was funded and undertaken or commissioned by military, aviation and aeronautical enterprises.

As mentioned earlier, Wierwille (1988) suggested that mental workload research made initial strides and then appeared to have become bogged down in details. This could be possible, but it may also be possible that researchers have moved on to new (and greener?) research pastures, such as fatigue management, stress and situational awareness. Twenty-eight years after the peak period in mental workload research, there is still no consensual definition of mental workload among researchers. This could have an affect on the interest and motivation of prospective researchers in choosing mental workload as an area of research.

Between 1990 and 2003 very little original work was produced and certainly no new tools for the measurement of mental workload saw the light. Publications from this period focus mainly on the investigation and refinement of physiological measurements of mental workload.

There has however been a resurgence of interest in the relevance of human factors in the railway environment, especially in terms of the mental workload of train drivers and train controllers or dispatchers. Several papers have been published since 2003 and experimental work with new mental workload assessment tools has been conducted. This research is reported on in paragraph 2.9.

A comprehensive listing of papers and presentations on the investigation of human mental workload has been compiled by Hancock, Mihaly, Rahimi and Meshkati (1988).

The literature on workload measurement between 1970 and 1990 reflects a trend towards mental workload quantification, which has resulted in the development of several measures, test instruments and analytical procedures, collectively called *workload estimation techniques* or *workload measurement methodologies*.

Designers and operators realise that performance is not all that matters in the design of a good system. It is just as important to consider what demands a task imposes on operators' limited resources. More specifically, the importance of research on mental workload may be viewed in three different contexts: workload prediction, the assessment of workload imposed by equipment, and the assessment of workload experienced by operators. The difference between the second and the third context exists in their implications for action. When the workload of systems is assessed or compared, the purpose of such a comparison is to optimise the system. When the workload experienced by operators is assessed, it is for the purpose of choosing between operators or providing operators with further training (Wickens and Hollands, 2000).

The assessment of workload, whether psychological or physical, commonly relies on the resource construct, meaning that there is a given (measurable) quantity of capability and attitude (or capacity) available, of which a certain percentage is demanded by the task at hand. If less capacity than is available is required, a reserve exists. Accordingly, workload is often defined as the portion of resource (i.e., of the maximal performance capacity) expended in performing a given task. Following on from this concept, it is obvious that any situation in which more is demanded from operators than can be given should be avoided. This is because the task performance will not be optimal and operators are likely to suffer, physically or psychologically, from the overload (Kroemer, Kroemer and Kroemer-Elbert, 1994).

An increasing proportion of work taxes the information processing capabilities of operators, rather than their physical capacity. This is due to progressively increased automation. It is the load placed upon such cognitive capabilities that mental workload assessment is designed to measure (Meshkati et al., 1995).

The multidimensional nature of workload reflects the interaction of the multiple elements of task demands and operator variables as they relate to workload. When attempting to measure mental workload the challenge is to assess which of the behavioural, physiological or psycho-physiological dimensions will change as a result of mental workload, what form these changes will take, and how to measure them.

The model provided by Meshkati (1988) (Figure 2.5) indicates that the effects of mental workload will be reflected in directly or indirectly observable behaviour, which can be measured by means of various techniques. His model makes provision for three groups of measurement techniques, namely physiological measures, subjective measures and performance measures.

In terms of the operational definition of mental workload, a measurement technique would have to reflect elements of the input load and the capacity of operators. Furthermore, it should enable users of the system to predict the workload at a specific train control centre.

2.7 CRITERIA FOR THE SELECTION OF MENTAL WORKLOAD ASSESSMENT TECHNIQUES

In order to determine the suitability of mental workload measurement techniques for application in research and practical environments, several researchers have developed criteria against which the methods could be measured. The newly developed Mental Workload Index (MWLI) will also be measured against some of these criteria. These criteria will eventually determine the acceptability and utility of the index.

The measures that can be used for the assessment of mental workload have different properties and these properties range from very general to very specific aspects. A general aspect is, for instance, the amount of equipment that is needed. A more specific, and from a scientific perspective, more important aspect is the validity of a measure. Does the measure reflect the concept of mental workload as intended, or does it reflect other concepts, e.g., physical workload?

The following criteria are referred to in the literature (O'Donnell and Eggemeier, 1986; Gawron, Schflett and Miller, 1989):

1. Sensitivity: Is the technique able to detect changes in task difficulty or demands? The index should be sensitive to changes in task difficulty or resource demand.
2. Diagnosticity: Does the technique have the ability to discern the type or cause of workload, or the ability to attribute it to an aspect or aspects of the operators' tasks? An index should not only indicate when workload varies but also the cause of such variation.
3. Selectivity/Validity: The index should be sensitive only to differences in cognitive demands and not to changes in factors such as physical load or emotional stress, which may be unrelated to mental workload or information-processing ability. Types of validity include face, construct, content and predictive (validity).
4. Intrusiveness: The degree to which a technique degrades ordinary or primary task performance is called intrusiveness. The disruption in ongoing task performance as a result of the application of the measurement technique is an undesirable property and should be minimised. It can pose serious problems in the application of a workload measurement technique.
5. Reliability: Reliability refers to the workload estimate that has to be reliable both within and across tests – it must consistently reflect the mental workload. Measures that have been developed in a laboratory setting do not have to indicate workload equally well in the field. Between applications much will depend upon the region of task performance. A measure sensitive only to low levels of workload will not be able to discriminate between levels within high demand situations.
6. Implementation requirements: This criterion refers to the practical constraints associated with the complexity of the measurement procedure and apparatus, such as the requirement of specific equipment or operator training. In field studies in particular, implementation requirements can become important. A workload measurement should be easily transferable from the laboratory environment to the field situation. Factors that contribute to making a technique cumbersome include instrumentation, analyst and operator training, and data recording and analysis.

7. Operator and user acceptance: The degree of approval of the technique by the operators or the perception of the validity and usefulness of the procedure is referred to as operator acceptance. Operators' opinions about a measurement technique, especially the use of self-reports, can affect the correctness and accuracy of the measure. In general, acceptance is higher if the technique is less intrusive or artificial, while the face validity of specific measurements may enhance operator acceptance.
8. Affordability: The affordability of the application, administration and analysis of the technique remains an important criterion in the research process. The balance between affordability and reliable results is an important consideration, but such a balance is not easily obtained.

All the criteria mentioned should be considered and if possible, satisfied in the choice or development of a mental workload measurement technique. For the purposes of this study, the stakeholder group, namely the users of the new measurement technique, operators (TCOs) and employee representative organisations, were consulted on which criteria should be included. The following criteria were considered especially important:
Sensitivity
Selectivity/Validity
Intrusiveness
Implementation requirements
Operator acceptance
Affordability

2.8 DETAILED DISCUSSION OF MENTAL WORKLOAD ASSESSMENT TECHNIQUES

Workload measurement methods can be regarded as a practical intervention that attempts to translate the theoretical concept of *workload* into practical methods for the measurement and evaluation of workload and its effects on system performance. The terms *workload estimation techniques* or *workload measurement methodologies* refer to the numerous measurements, test instruments and analytical procedures that attempt to measure workload.

Mental workload per se cannot be directly observed and therefore its presence and its severity must be inferred from changes in overt behaviour and/or measurable physiological and psycho-physiological functioning (Casali and Wierwille, 1984).

The measurement techniques or methodologies in the literature therefore assume some link between workload and specific behavioural changes in the form of psycho-physiological or physiological change or change in performance. The basic paradigm is therefore that the extent and the nature of these changes occur as a direct consequence of changes in workload.

As far back as 1979, authors have developed different classifications for workload assessment methodologies, some more comprehensive than others (Williges and Wierwille; 1979; Johannsen, 1979; Jex, 1988; Eggemeier, 1988). Essentially these classifications are similar and only semantic differences are evident. For the purposes of this study, the different approaches to the assessment of mental workload are discussed under the following headings:

- Measures of primary task performance
- Secondary measures of spare mental capacity
- Subjective rating techniques
- Physiological indicators of mental workload

Johannsen (1979) added a measurement to his categorisation that was used effectively in the current study, namely Timeline Analysis (TLA). This method is discussed in Chapter 3.

Each of the approaches comprise several techniques, which are discussed under each heading (Kroemer et al., 1994; Meshkati et al., 1995; Wickens and Hollands, 2000).

Some of these assessment techniques meet many of the criteria proposed in the literature, but none satisfy all of the criteria. Some of the criteria may trade off with one another, and until now no technique has been found that satisfies all of the criteria.

2.8.1 Measures of Primary-Task Performance

Performance-based procedures, which include primary- and secondary-task measures, are based on operators' performance levels. Measures of primary-task performance use the adequacy of performance on the task to characterise capacity expenditure. Two types of performance measure have been identified: single measures of the primary task and multiple measures of the primary task. These measures attribute changes in the primary-task performance that can be measured to changes in the workload imposed by the task.

There are several methodological approaches to the measurement of performance, otherwise known as system-output measures (Meshkati et al., 1995).

The analytical approach looks in detail at the actual performance of the task to be assessed, examining not only overall achievement but also the way in which it was attained. From a practical perspective the analytical approach appears most appropriate. The advantage of this method is that the various decisions and other processes that constitute performance are considered in the context in which they normally occur, so that the complexities of any interaction between different elements in the task can be observed.

This approach presents two difficulties. Firstly, the detailed scores required may be difficult to obtain for tasks such as process monitoring, in which most of the decisions made do not result in any overt action. Secondly, even where there is sufficient observable action, recording may have to be elaborate, and analysis of the results can prove laborious.

Synthetic methods provide another approach to performance measurement as a mental workload assessment technique. Specific performance demands that may be placed on operators are identified through task analysis. Performance times and operators' reliability are assigned to the distinct tasks. Information on performance time is then accumulated for a given phase, and the total is compared with the predicted duration of the phase. This comparison of required time with available time can be employed as an index of workload.

Benefits and Shortcomings of Primary-Task Measures

Primary-task measures are probably the most obvious method of mental workload assessment, and one would assume that this method would satisfy a number of the criteria for the selection of techniques, such as ease of implementation and user acceptance. However, the shortcomings far outweigh the benefits of this approach.

There are four reasons why the primary-task performance approach may prove inadequate in clearly revealing the measures of the primary task:

- Firstly, two primary tasks may lie in the *underload* region of the supply-demand space (see Figure 2.1). Since both have sufficient reserve to reach perfect performance, this measure cannot discriminate between them.
- Secondly, there may be differences in how two primary tasks that are to be compared are measured, and what those measures mean.
- Thirdly, sometimes it is simply impossible to obtain good measures of primary-task performance (as with measures of decision-making or vigilance tasks).
- Finally, two primary tasks may differ in their performance, not by the resources demanded to achieve that performance, but by differences in data limits.
  (Wickens and Hollands, 2000).

The lack of sensitivity of performance measures to changes in mental workload levels is one of the major problems of these methods. The level of mental workload may increase while performance is unchanged, so that performance may not be a valid measure of workload. The generalised application of this method to different task situations poses another set of problems, since for each experimental situation, a unique measure must be developed. This problem is not shared by the other methods of mental workload assessment (Meshkati et al., 1995).

Numerous studies support the same conclusion, namely that no substantial change occurs in the primary task as a function of workload (Williges and Wierwille, 1979).

Further, unless one is willing to assume that humans always operate to capacity and that all humans have the same capacity, the performance-based workload scales would reflect only the states of the particular individuals from whom the data were collected. In other words, inter-individual comparisons may not be valid for this measure (Meshkati et al., 1995).

Other problems with primary-task measures are intrusiveness and the interpretation of results. Primary-task data may reflect a wide variety of influences, such as motivation and learning effects (Meshkati et al., 1995).

Williges and Wierwille (1979) concluded that only high workload situations (those that approach operator overload) are discernable by primary-task performance measures, while low workload conditions may not be because at these low levels operators ordinarily adapt in an effort to maintain output variables at an acceptable level.

For these reasons, system designers have often turned to the three other workload assessment techniques, which may assess more directly either the effort invested in primary-task performance or the level of residual capacity available during that performance.

2.8.2 Secondary Measures of Spare Mental Capacity

The largest body of research that deals with operator mental workload is focused on the evaluation of the concept of spare (residual or reserve) mental capacity. This concept is founded on the assumption of a limited-channel model of the human operator (Williges and Wierwille, 1979). Secondary-task measures are typically derived from the levels of performance on a concurrent or secondary task (Eggemeier, 1988).

According to the above-mentioned authors, spare mental capacity is the difference between the total workload capacity of operators and the capacity needed to perform the task. As spare mental capacity decreases, operators' workload increases until the point of overload is reached. At this point, the information processing demands of the task exceed the workload capacity of the operators. This approach assumes that an upper limit exists on the ability of human operators to gather and process information.

A study of the research literature rendered three approaches in the measurement of spare mental capacity: task analytic methods, secondary-task measures, and occlusion procedures.

(a) Task Analytic Methods

Task analytic methods are based on mathematical methods derived from systems engineering. This approach assumes that all task components, when performed serially, require specific lengths of time to complete, and as long as the time available for overall completion exceeds the sum of theoretical time durations for performing the task components, spare mental capacity exists. When, however, the actual time available is insufficient, stress and task loading occur (Wierwille and Gutmann, 1978).

The methods used under the Task Analytic approach are:

- *Task Component/Task Summation*: System engineering principles and task analysis procedures have been widely used to develop computer-based models of aircrew and cockpit simulation techniques with which to obtain descriptions of aircrew performance. An essential input is the detailed task analysis that forms the basis of time assessments for each task component.
- *Information Theoretic*: This approach attempts to quantify workload in terms of an information-transmission metric that is specified in bits/second. This approach has been applied in the evaluation of visual monitoring but has not found wide application and support (Wierwille and Gutmann, 1978).

(b) Secondary-Task Methods

Secondary tasks have been used in most behavioural research approaches that seek to estimate spare mental capacity. A secondary task is a task that operators are asked to perform in addition to their primary task. The task is performed only once the primary task has been fully attended to. This method is therefore an indirect measure of operator workload. If operators are able to perform well on the secondary task, this indicates that the primary task is relatively easy; if they are unable to perform the secondary task and simultaneously maintain the primary task performance, this indicates that the primary task is more demanding. The difference between the performances obtained under the two conditions is then taken as a measure or index of the workload imposed by the primary task (Meshkati et al., 1995).

Measuring performance on a secondary concurrent task is intended to assess the spare capacity that remains after capacity resources have been allocated to the primary task. If subjects allocate some of the resources actually required for primary-task performance to the secondary task, the secondary task is intrusive (or invasive) on the primary task. An intrusive secondary task modifies the condition that is to be assessed.

A multitude of secondary tasks have been proposed and employed over the last two decades to assess the residual capacity available after the completion of primary tasks. Some of the most prominent examples are described here (Wickens and Hollands, 2000):

- Rhythmic tapping task: Operators must produce finger or foot taps at a constant rate. Tapping variability increases as primary workload increases.
- Random number generation: This requires that operators generate a series of random numbers. Normally the degree of randomness declines as workload increases, and more repetitive sequences are generated.
- Probe reaction time: These tasks are often used as a workload measurement technique, as it is assumed that greater primary-task workload will prolong the reaction time to a secondary-task stimulus.
- Other techniques: These include arithmetic addition, and critical tracking tasks.
- Handwriting analysis: Handwriting deteriorates as a result of distraction that can result from having to perform other tasks. This deterioration is a potentially effective measure of mental overload.

(c) Occlusion procedures

Occlusion is a time-sharing task that is forced rather than voluntary. Operators are given samples of the visual information that is required to perform the primary task and then information inputs, i.e. from the visual display, are suppressed or blocked. This can be achieved by operators having to wear a helmet fitted with an opaque visor that can be closed by external control. Alternatively, visual input is blocked by directly blanking the electronic displays of the system. This method is however considered intrusive and potentially unsafe (Meshkati et al., 1995).

(d) Benefits and Shortcomings of Secondary-Task Measures

There are two distinct benefits of secondary-task techniques. Firstly, they have a high degree of face validity. They are designed to predict the amount of residual attention operators will have available if an unexpected failure or environmental event occurs. Secondly, the same secondary task can be applied to two very different primary tasks and it will give workload measures in the same units (Wickens and Hollands, 2000).

However, the secondary task as a mental measurement technique has many shortcomings. Perhaps the most difficult aspect of secondary-task methodology for assessing workload is intrusiveness. When the secondary task is introduced, it normally interferes with the primary task and performance on the primary task is known to be modified and usually degraded. This technique therefore suffers on the intrusiveness criterion (Williges and Wierwille, 1979; Wierwille, Rahimi, and Casali, 1985).

One the one hand, this may be inconvenient or even dangerous if the primary task is one like flying, driving or controlling the movement of trains, and a diversion of resources to the secondary task at the wrong time could lead to an accident. On the other hand, disruption of the primary task could present problems of interpretation if the amount of disruption suffered by two primary tasks to be compared is not the same, that is, the measurement technique differentially disrupts that which is being measured (Wickens and Hollands, 2000).

2.8.3 Subjective Rating Techniques

Subjective assessment or self-reporting is probably the most common method used to evaluate mental workload. This method provides a valid assessment of the overall workload inflicted on the working memory of operators. The ease and speed of application and interpretation, as well as the high face validity of these measurements make them popular to use (Vidulich, 1988).

Although various individual assessment techniques have been developed, all subjective procedures use some report of experienced effort or capacity expenditure to characterise workload levels. Subjective methods include direct or indirect questioning of individuals for their opinions of the workload involved in tasks (Eggemeier, 1988; Meshkati et al., 1995).

Subjective measures represent the operators' conscious judgements regarding the difficulties encountered in the performance of the evaluated task. If a task imposes a high workload on operators, then it is expected that operators will *feel* loaded and will be able to report it. Care should however be taken with this approach. It is essential to understand which factors determine the relationship between subjective workload and objective performance. A subjective workload assessment should be viewed essentially as a verbal report on the level of information load experienced when performing a task (Vidulich, 1988).

Subjective assessments of the perceived workload rely on internal integration of the demands, but they may be unreliable, invalid, or inconsistent with other performance measures. If subjective measures are taken after the task has been completed, they are not real-time evaluations; if performed during the task, they may intrude (Nygren, 1991; Wickens and Yeh, 1983).

Some proponents of self-report measures consider them to be the best measures since they come nearest to tapping the essence of mental workload. Critics, on the other hand, say that the source of resource demands is hard to introspectively diagnose within a dimensional framework. Physical and mental workload is, according to the critics, hard to separate (O'Donnell and Eggemeier, 1986). For reasons that are discussed later in this chapter, self-report measures were not considered viable for the Spoornet study.

Pulat (1992) expresses his view of subjective ratings as follows: "The easiest but probably the most unreliable method is simply to ask operators their opinion on their workload on completion of a task or shift." He provides an example of a list of key words or definitions describing the different levels of workload that are typically used in rating scales:

Task difficulty	Low .... High
Time pressure	None .... Rushed
Effort	None .... Impossible
Frustration	Fulfilled .... Exasperated
Stress level	Relaxed .... Tense
Fatigue	Alert .... Exhausted

Another approach to the subjective rating of workload is to use interviews and questionnaires. The procedures used in this approach are not as structured as rating scales. They range from completely open-ended debriefing sessions to self-reporting logs of stressful activities (Meshkati et al., 1995).

Subjective measures of workload, such as the modified Cooper-Harper Scale (Wierwille et al., 1985), the NASA Task Load Index (TLX) (Hart and Staveland, 1988) and the Subjective Workload Assessment Technique (SWAT) (Reid and Nygren, 1988) have been widely used. According to a personal communication from a human factors expert in the transport industry in the USA, the SWAT and NASA-TLX are considered to be industry standards and therefore a good benchmark when developing a measurement tool (Reinach, 2006a). These instruments have, however, been criticised on both theoretical and technical grounds.

An individual's perceived difficulty (of a task) is influenced by long-term memory, which includes both memory of general experience and memory of similar tasks. Background factors, such as personality traits, habits and general attitudes as well as transient conditions (e.g. emotional state, fatigue, motivation and anticipated success or failure) play a role in operators' subjective ratings of their experience of the task (Meshkati et al., 1995).

Care must be taken when interpreting these scales since they contain words that may be interpreted differently under different task situations.

The rating scales and questionnaires identified in the research literature are discussed in more detail below.

(a) Variations of the Cooper Scale

Since the 1960s the mental load imposed by manual control tasks, specifically aircraft handling qualities, has been measured by the Cooper (C) Scale and modifications thereof, namely the Cooper-Harper (C-H) and Modified Cooper Harper (MC-H) scales (Moray, 1982).

The Cooper Scale (Moray, 1982) offers ten statements to the evaluator, who has to indicate which statement best approximates his opinion of the handling qualities of the aircraft or aspect under consideration. Various authors have commented on the deficiencies of this instrument. Criticism of the instrument relates to: (i) wording ambiguity and confusing nomenclature; (ii) the dual mission character of the scales and mixing of the tasks, e.g., normal and emergency conditions; and (iii) the lack of information about the quantitative character of the psychological scale continuum (Cilliers, 1992).

Table 2.1: The Original Cooper Rating Scale (reproduced in Cilliers, 1992.)

	Adjective Rating	Numerical Rating	Description	Primary mission accomplished	Can be landed
NORMAL OPERATION	Satisfactory	1	• Excellent, includes optimum	Yes	Yes
		2	• Good, pleasant to fly	Yes	Yes
		3	• Satisfactory, but with some mildly unpleasant characteristics	Yes	Yes
EMERGENCY OPERATION	Unsatisfactory	4	• Acceptable, but with unpleasant characteristics	Yes	Yes
		5	• Unacceptable for normal operation	Doubtful	Yes
		6	• Acceptable for emergency condition only	Doubtful	Yes
NO OPERATION	Unacceptable	7	• Unacceptable, even for emergency condition	No	Doubtful
		8	• Unacceptable – Dangerous	No	No
		9	• Unacceptable - Uncontrollable	No	No
	Unprintable	10	• “Motions possibly violent enough to prevent pilot escape”		

The Cooper-Harper (C-H) Scale is the most widely used rating scale for assessing the handling qualities of an aircraft. The descriptors of the scale pertain to the *flyability* of an aircraft, and although the scale contains some reference to workload, the descriptors would have to be modified for use in other workload applications (Cilliers, 1992).

The C-H Scale is one of the more validated scales for the subjective measurement of workload on aircraft handling qualities. It consists of a 10-point scale with a decision-tree format. It makes provision for performing the rating task sequentially, reaching a final rating in a deliberate and careful manner (Cilliers, 1992).

The main disadvantage of decision-tree rating scales is that they cannot provide information of an equal interval nature. At best they provide ordinal estimates of workload (Jex, 1988).

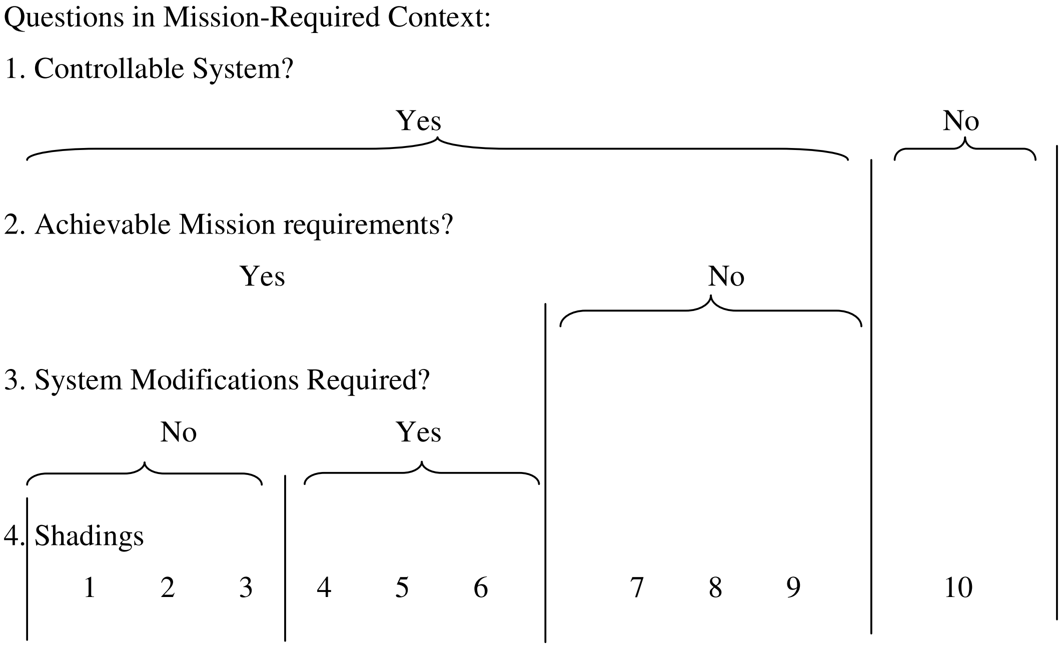
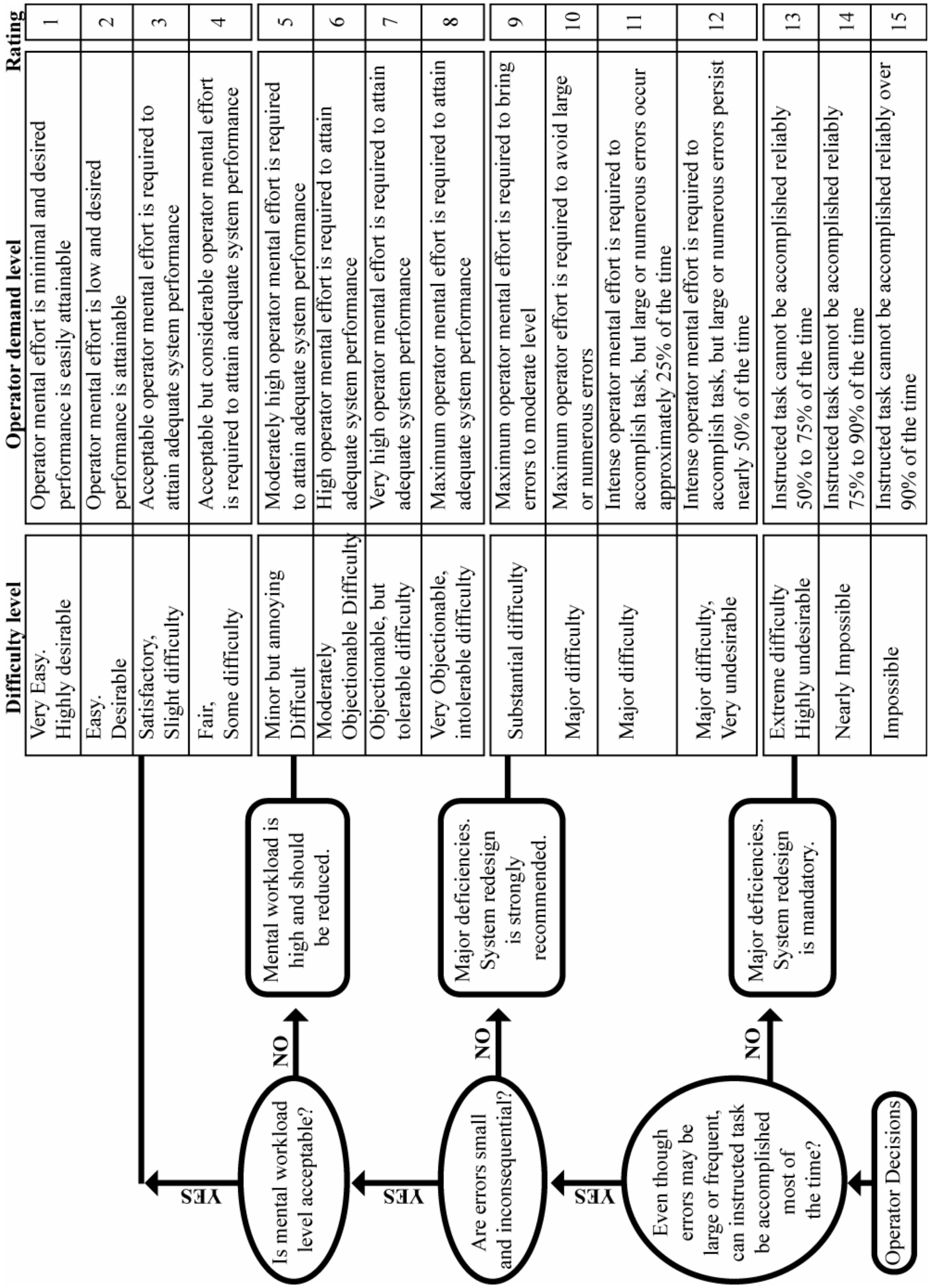


Figure 2.6: Cooper-Harper Rating Scale (reproduced in Cilliers, 1992)

The Modified Cooper-Harper (MC-H) Scale is considered a further development in the subjective measurement of mental workload. This scale is applicable to a wider variety of task workloads, especially for systems that may load perceptual, mediational and communication activities. The scale can be applied to obtain mental workload estimations regardless of the type of loading imposed by the task. The wording of the scale has been modified to enable the assessment of mental workload as distinct from the psychomotor workload, which the original scale was designed to measure (Rahimi and Wierwille, 1982; Wierwille and Casali, 1983a, Wierwille et al., 1985).

Figure 2.7 provides an excerpt from the Modified Cooper-Harper Scale.



Difficulty level	Operator demand level	Rating
Very Easy. Highly desirable	Operator mental effort is minimal and desired performance is easily attainable	1
Easy. Desirable	Operator mental effort is low and desired performance is attainable	2
Satisfactory, Slight difficulty	Acceptable operator mental effort is required to attain adequate system performance	3
Fair, Some difficulty	Acceptable but considerable operator mental effort is required to attain adequate system performance	4
Minor but annoying Difficult	Moderately high operator mental effort is required to attain adequate system performance	5
Moderately Objectionable Difficulty	High operator mental effort is required to attain adequate system performance	6
Objectionable, but tolerable difficulty	Very high operator mental effort is required to attain adequate system performance	7
Very Objectionable, intolerable difficulty	Maximum operator mental effort is required to attain adequate system performance	8
Substantial difficulty	Maximum operator mental effort is required to bring errors to moderate level	9
Major difficulty	Maximum operator effort is required to avoid large or numerous errors	10
Major difficulty	Intense operator mental effort is required to accomplish task, but large or numerous errors occur approximately 25% of the time	11
Major difficulty, Very undesirable	Intense operator mental effort is required to accomplish task, but large or numerous errors persist nearly 50% of the time	12
Extreme difficulty Highly undesirable	Instructed task cannot be accomplished reliably 50% to 75% of the time	13
Nearly Impossible	Instructed task cannot be accomplished reliably 75% to 90% of the time	14
Impossible	Instructed task cannot be accomplished reliably over 90% of the time	15

Figure 2.7: Modified Cooper-Harper Scale and rating scales (reproduced in Cilliers, 1992)

(b) The Bedford Scale

The Bedford Scale is described as similar to the Modified Cooper-Harper Scale, in which a single score is generated on a 10-point scale. A workload of 10 is considered to be extremely high with no spare capacity and a level of 1 is considered to be insignificant.

The underlying assumption of the Bedford Scale is that workload is best defined in terms of the subjective experience of *effort*. This definition takes into account moderating factors such as ability, experience and the individual’s response to stress.

Critics of the Bedford Scale are of the opinion that the scale is not linear and lacks sensitivity at the lower end. The scale has been used in field studies but has not undergone laboratory validation (Cilliers, 1992).

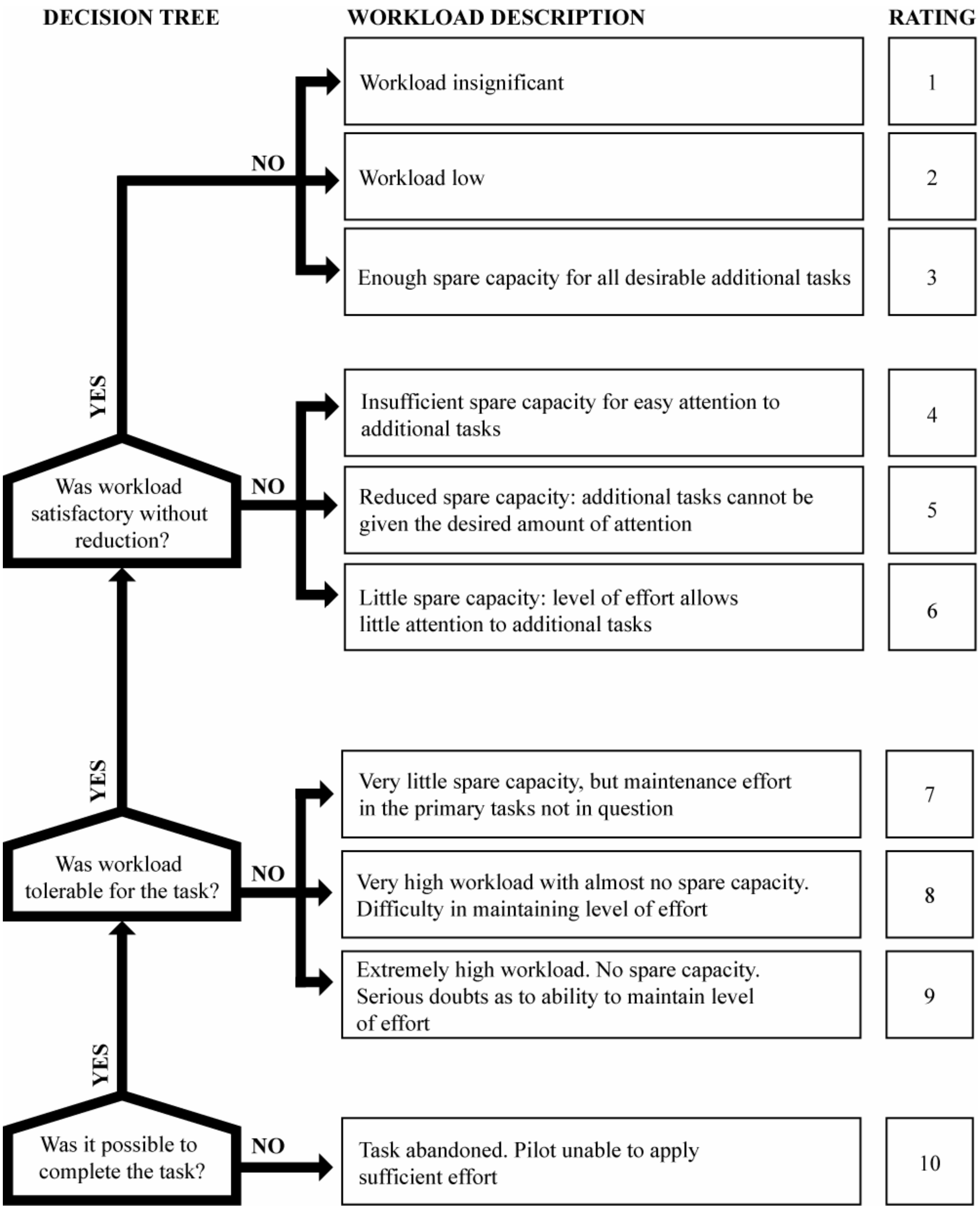


Figure 2.8: The Bedford Scale (reproduced in Cilliers, 1992)

(c) The Subjective Workload Assessment Technique (SWAT)

Another widely used subjective rating technique is the Subjective Workload Assessment Technique, which is a multidimensional rating scale, specifically designed to measure operator workload in a variety of systems, and for a number of tasks (Reid and Nygren, 1988, Meshkati et al., 1995).

The SWAT has been developed as a generalised procedure for scaling pilot mental workload. The procedure entails a two-step process wherein each subject completes a scale development phase and an event-scoring phase. During the scale development phase, data necessary to develop a workload scale is obtained from a group of subjects. During the event-scoring phase, subjects rate the workload associated with a particular task and/or mission segment. These two steps are considered to be distinct events which occur at different times (Reid and Nygren, 1988).

The principles that guided the development of SWAT were the following:

- To develop as precise a measure as possible while minimising the intrusiveness of the data collection procedure in the operational situation;
- To place minimal measurement constraints on the complexity of the judgement task that is required of the operators making workload evaluations;
- To provide a mechanism for testing the validity of the formal measurement model that is assumed by the underlying additive model in SWAT (Reid and Nygren, 1988).

The SWAT measures workload on three, three-point scales (See Table 2.2).

Table 2.2: The SWAT Scale (From: Wickens and Hollands, 2000)

Time Load	Mental Effort Load	Stress Load
1. Often have spare time. Interruptions or overlaps among activities occur infrequently or not at all.	1. Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.	1. Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.
2. Occasionally have spare time. Interruptions or overlaps among activities occur frequently.	2. Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability or unfamiliarity. Considerable attention required.	2. Moderate stress due to confusion, frustration, or anxiety noticeably adds to the workload. Significant compensation is required to maintain adequate performance.
3. Almost never have spare time. Interruptions or overlaps among activities are very frequent or occur all the time.	3. Extensive mental effort and concentration necessary. Very complex activity requiring total attention.	3. High to very intense stress due to confusion, frustration or anxiety. High to extreme determination and self-control required.

Although SWAT is one of the most sophisticated workload assessment techniques available, it is by no means free of problems. The following are unresolved practical and theoretical issues (Hart, 1986):

- The between-subject variance is high and standard deviations can be as high as 70-80% when compared with average SWAT scores.
- The scale development phase is time consuming.
- The assumption that people can accurately predict the workload of 27 combinations of abstract variables is probably optimistic and not justifiable.
- High inter-rater reliability may only reflect agreement on extremes of workload.

- Three factors alone are probably not sufficient to characterise workload experiences and definitions.

SWAT demonstrated sensitivity for workload assessment related to visual display monitoring (the scales were designed to heavily load perceptual input capacity); verbal and spatial short-term memory (the scales primarily loaded two major central processing coding dimensions); and unstable tracking (the scales exerted heavy demands on motor output capacity). SWAT scales (like the M C-H and NASA-TLX Scales) are capable of reflecting variations in effort expenditure across a variety of processing functions, but should be considered as global measures of workload rather than diagnostic in their sensitivity (Eggemeier, 1988).

(d) The NASA Task Load Index (TLX)

The work of Hart and Staveland culminated in the development of the NASA Task Load Index (TLX), an extensively researched and widely used subjective mental workload measurement tool that is considered one of the gold standards in the subjective assessment of mental workload.

The NASA-TLX was developed as part of a research program that was conducted over several years to identify the factors associated with variations in subjective workload within and between different types of tasks. The tasks included simple cognitive and manual control tasks, complex laboratory and supervisory control tasks and aircraft simulation (Hart and Staveland, 1988).

A multidimensional rating scale was proposed by Hart and Staveland, in which information about the magnitude and source of six workload-related factors were combined to derive a sensitive and reliable estimate of workload.

Table 2.3: The NASA Bipolar Rating Scale descriptions (Hart and Staveland, 1988)

RATING SCALE DESCRIPTIONS		
Title	Endpoints	Descriptions
OVERALL WORKLOAD	*Low, High*	The total workload associated with the task, considering all sources and components.
TASK DIFFICULTY	*Low, High*	Whether the task was easy or demanding, simple or complex, exacting or forgiving.
TIME PRESSURE	*None, Rushed*	The amount of pressure you felt due to the rate at which the task elements occurred. Was the task slow and leisurely or rapid and frantic?
PERFORMANCE	*Failure, Perfect*	How successful you think you were in doing what we asked you to do and how satisfied you were with what you accomplished.
MENTAL/ SENSORY EFFORT	*None, Impossible*	The amount of mental and/or perceptual activity that was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)?
PHYSICAL EFFORT	*None, Impossible*	The amount of physical activity that was required (e.g., pushing, pulling, turning, controlling, activating, etc.).
FRUSTRATION LEVEL	*Fulfilled, Exasperated*	How insecure, discouraged, irritated, and annoyed versus secure, gratified, content, and complacent you felt.
STRESS LEVEL	*Relaxed, Tense*	How anxious, worried, uptight, and harassed or calm, tranquil, placid and relaxed you felt.
FATIGUE	*Exhausted, Alert*	How tired, weary, worn out, and exhausted or fresh, vigorous, and energetic you felt.
ACTIVITY TYPE	*Skill-based, Rule-based, Knowledge-based*	The degree to which the task required mindless reaction to well-learned routines or required the application of known rules or required problem solving and decision making.

Each scale had bipolar descriptors at each end (high/low). Numerical values were not displayed, but values ranging from 1 to 100 were assigned to scale positions during data analysis. This set of scales was used to evaluate the experiences of subjects in 25 different studies.

One of the goals in the gathering of workload and workload ratings was to amass a database that would allow for the examination of the relationships among different task, behaviour and psychological factors in order to create a valid and sensitive rating technique for subjective workload assessment. The assumption was made that the technique would be multidimensional, but that the number of subscales should be less than the number used for research purposes.

In the comprehensive experimentation that followed, data was collected from various types of tasks: single cognitive tasks, single axis manual control tasks, dual task experiments and various simulated tasks. A comprehensive amount of data was collected and analysed to construct a workload rating scale.

From their experimentation Hart and Staveland (1988) reported a number of key points that emerged from the subjective experiences and evaluations of workload by the subjects:

- A phenomenon exists that can generally be termed *workload*, but its specific causes may differ from one task to the next.
- Ratings of component factors are more diagnostic than overall workload ratings.
- Subjects' workload definitions are different, which contributes to between-subject variability. However, the specific sources of loading imposed by a task are more potent determinants of workload experiences than such *a priori* biases.
- A weighted combination of the magnitudes of factors that contribute to subjects' workload experience during different tasks provides an integrated measure of overall workload that is relatively stable between raters.

The research resulted in the development of the NASA-TLX rating scale of which the following sub-scales form a part and which and are rated by the subjects:

Table 2.4: The NASA-TLX Rating Scale Definitions (Hart and Staveland, 1988).

NASA-TLX RATING SCALE DEFINITIONS		
Title	Endpoints	Descriptions
MENTAL DEMAND	*Low, High*	How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
PHYSICAL DEMAND	*Low, High*	How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
TEMPORAL DEMAND	*Low, High*	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
PERFORMANCE	*Good, Poor*	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
EFFORT	*Low, High*	How hard did you have to work (mentally and physically) to accomplish your level of performance?
FRUSTRATION LEVEL	*Low, High*	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

This scale has been subjected to extensive studies of reliability and validity (Hart and Staveland, 1988). However, none of these studies involved the physiological measurement of the stress parameters associated with mental workload.

(e) The Workload Profile (WP)

More recently the Workload Profile (WP), a multidimensional instrument for the assessment of subjective mental workload, was developed and introduced. This instrument is based on the multiple resource model of Wickens (1992). The WP attempts to combine the advantages of secondary-task performance-based procedures (high diagnostic ability) and subjective techniques (high subject acceptability and low implementation requirements and intrusiveness). The WP asks subjects to provide the proportion of attention resources that they had used after they had experienced all of the tasks.

In a recent study the characteristics of the WP were compared with NASA-TLX and SWAT. The WP bore the highest sensitivity (the tool's ability to detect changes in task difficulty demands). It also showed high diagnostic ability and low intrusiveness. Problems with implementation and acceptability were however experienced (Rubio, Diaz, Martin and Puente, 2004).

(f) Benefits and Shortcomings of Subjective Techniques

The benefits of subjective techniques are apparent. They do not disrupt primary-task performance, and they are relatively easy to apply and interpret. Their shortcomings relate to the uncertainty with which operators' verbal statements diagnostically reflect the investment of or demand for processing resources and are not influenced by other biases (e.g., dislike of or unfamiliarity with the task, or the rater's reluctance to report that tasks are difficult).

Among the shortcomings of subjective ratings is the problem posed by a task being primarily a function of the raters' perceptions. The perceived difficulty of a task might alter operators' attitudes to it. Subjective ratings of task difficulty could also be affected by the situation and the job as a whole rather than by task-induced or individual rater factors alone (Meshkati et al., 1995).

Furthermore, these methods are insensitive to demands outside the component of the human information-processing system that deals with working memory. Also, performing multiple tasks concurrently (as is the case with TCOs in this project) seems to render subjective assessments somewhat insensitive to changes in one or more of the tasks.

Air traffic controller (ATC) workload evaluations have tended to rely mainly on subjective measures. Although the use of subjective measures is attractive (they are inexpensive and easily collected), researchers should remain mindful of their limitations. This admonition seems especially relevant to the ATC domain, in which highly skilled operators are asked to give subjective reports. Such operators may be unable – or for reasons of job security, unwilling – to give accurate reports. Moreover, such reports may be subject to individual biases, preconceptions, or memory limitations (Hilburn and Jorna, 2001).

2.8.4 Physiological Indicators of Mental Workload

One solution to performance intrusiveness is to record, unobtrusively, the manifestations of workload through appropriately chosen physiological measures of autonomic or central nervous system activity. These techniques offer two distinct advantages over secondary-task and subjective measures: participants are not required to execute any extra overt behaviour, and psycho-physiological measures are capable of measuring fluctuations in mental workload over time.

The advantage of physiological responses measurements is that they do not require overt responses from operators, and most cognitive tasks do not require overt behaviour. Moreover, most of the measures can be collected continuously, and measurements have now become relatively unobtrusive due to the miniaturisation of instruments (De Waard, 1996).

Human mental workload requires the expenditure of physiological effort and resources. The group of workload techniques termed *physiological* can generally be discerned from other workload estimation techniques by the fact that the changes that cause measurable variations are largely involuntary. The autonomic (involuntary) nervous system is of particular interest in the context of workload measurement as it controls the heart, secretion glands and the involuntary muscles (Wilson and O'Donnell, 1988).

Physiological measures reflect processes such as the demand for increased energy, progressive degradation of the system or homeostatic functioning of mechanisms designed to restore system equilibrium that have become disturbed by cognitive requirements. It is, however, difficult to distinguish between activation that is specific to the perceptions of individual operators, and the activation that results from the actual workload imposed (Hancock, Meshkati, Robertson and Robertson, 1985). In any situation in which individuals are required to increase mental or physical effort or to deal with stimuli that evoke emotion, involuntary changes in the autonomic nervous system (ANS) are produced. Although it would be reasonable to assume that these changes would be reflected in physiological measurements, attempts at finding such a specific mental workload index have been unsuccessful.

2.8.4.1 Mental Workload and Stress Measurement

To understand how mental overload manifests itself physiologically and why physiological measurements are considered useful in measuring workload, a discussion on the functional relationship between workload and stress is necessary.

Physiological measures are one of the most widely used research methods for assessing operator mental workload. The physiological method generally involves measuring and processing data for one or more variables related to human physiological processes. The underlying concept in physiological monitoring is as follows: As operator workload increases or decreases, involuntary changes take place in the physiological processes of the human body (body chemistry, nervous system activity, circulatory or respiratory activity). Consequently, workload may be assessed by measuring and processing the appropriate physiological variables.

Among many researchers there is an underlying assumption that high workload levels are accompanied by increased emotional stress. This stress is then measured by physiological recordings and is related back to workload. Stress in this case is assumed to act as an intermediate variable, causing physiological changes. Among other researchers the underlying assumption involves changes in the state of arousal. *Arousal* may be considered to be a state of preparedness or level of activation of the human organism (Wierwille, 1979).

Through changes in the status of a number of physiological systems, individuals engaged in cognitive activities provide indirect indices of their level of effort. Ursin and Ursin (1979) recognised that these physiological methods do not measure the imposed load but rather give information as to how individuals respond to the load and, in particular, whether they are able to cope with it.

Successful physiological measurement of operator workload has, unfortunately, been much easier to conceptualise than to achieve. Workload would intuitively seem to require the expenditure of physiological effort and resources, and it is reasonable to assume that some central or peripheral measure could be found that would provide an index of this expenditure. Early attempts to find such an index were, however, remarkably unsuccessful. The frequent failure of specific physiological measures to correlate with imposed workload has been noted (O'Donnell and Eggemeier, 1986). Indeed such failures led to an early belief that physiological measures might have questionable value as valid, reliable measures of mental workload. However, theoretical and laboratory work continued to include and refine a variety of such techniques, such as heart rate, eye blink and EEG measures. These efforts met with mixed success, and while the inconclusive results tended to stimulate interest in physiological measures, the preponderance of data continued to be negative, ambivalent or contradictory (Wilson and O'Donnell, 1988).

It should be recognised that many aspects of operator behaviour other than mental workload may have an effect on physiological measures. A large number of researchers who did not obtain significant results in applying mental workload measurement techniques suggested that either the sample population must be homogenised, or that personality traits, individual differences and other related factors such as past experience, skill, and motivation should be incorporated in the model, since many of these factors have been shown to influence physiological measurements (Meshkati and Loewenthal, 1988).

Early indications of specificity for at least some physiological measures came from studies such as those by Lacey and Lacey (1958). These investigators demonstrated the remarkable ability of cardiac measures to differentiate between various types of task-related ECG deceleration patterns. Various aspects of the cardiac cycle were shown to be dependent on both the task activity and the mental and/or physical involvement of the individual.

It is recognised that the functional relationship between mental workload and stress is not a very firm one, due to the complexity of the concepts of mental workload and stress and the variety of other factors that influence the measurement of these concepts. Also, one single measure has proven insufficient to measure these concepts. Despite these difficulties, it is the opinion expressed in this paper that physiological measurements still provide a valuable objective measure of mental workload. The next section explains the relationship between mental load and the effect on the human nervous system.

2.8.4.2 The Organisation of the Human Nervous System

In order to understand how mental overload manifests itself physiologically and why physiological measurements are considered useful in the measurement of workload, an understanding of physiological stress manifestation mechanisms is necessary.

From an anatomical perspective, two fundamental nervous systems exist: the central nervous system (CNS) and the peripheral nervous system (PNS):

(a) CNS

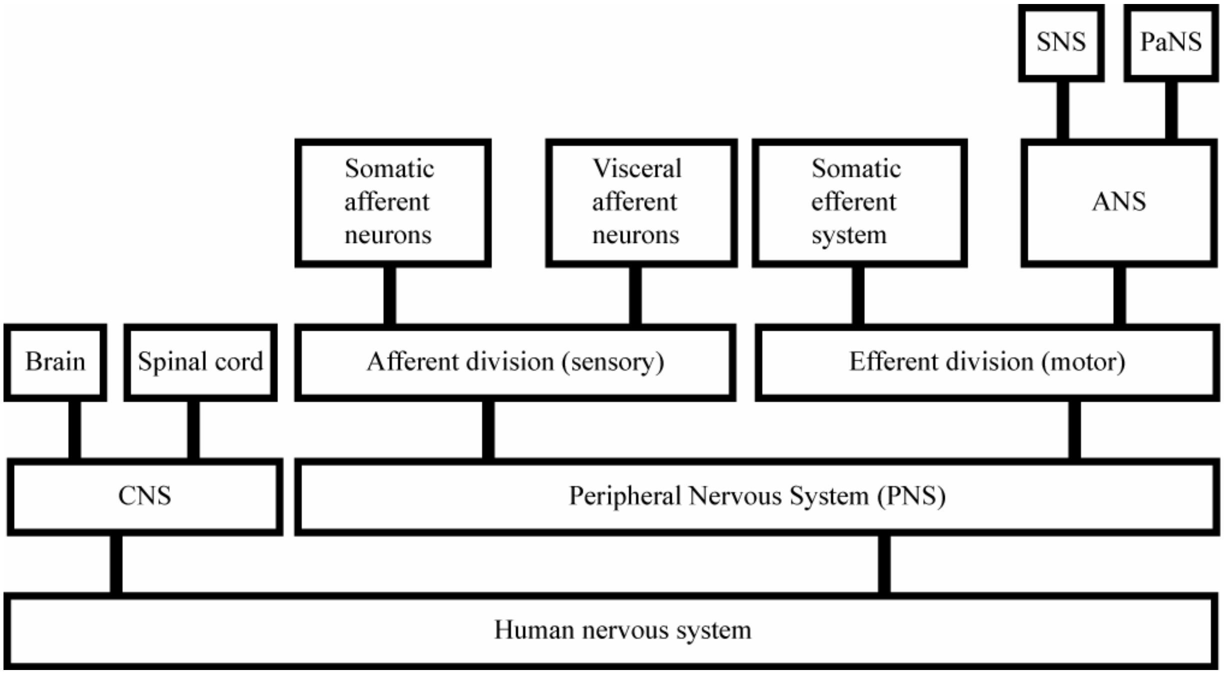
- The CNS consists of the brain and spinal cord. Together they function as the integrated control centre of the complete nervous system. Incoming sensory information is serviced by responses based on experience (memory), reflexes and current conditions (Gericke, 1994).
- The CNS is further divided into three functional levels (Everly and Rosenfeld, 1981):
    - The **neocortex** is the most sophisticated component of the human brain. Among other functions such as the decoding and interpretation of sensory signals, communication, and gross control of motor behaviour (musculoskeletal), the neocortex (primarily the frontal lobe) presides over imagination, logic, decision making, memory, problem solving, planning and apprehension.

  - The **limbic system** (or limbic brain) represents the major component of the second level of the brain. The limbic brain is of interest in the discussion of stress because of its role as the emotional (affective) control centre for the human brain. The limbic system consists of numerous neural structures, e.g., the hypothalamus, hippocampus, septum, cingulated gyrus, amygdala and pituitary gland (the master endocrine gland).
  - The **reticular formation** and the **brain stem** represent the third and most primitive level of the brain. The major function of this level is the maintenance of vegetative functions (heartbeat, respiration, vasomotor activity) and the conduction of impulses through the reticular formation and relay centres of the thalamus to the higher order levels of the brain.
  - The **spinal cord** represents the central pathway for neurons as they conduct signals to and from the brain.

(b) PNS

The PNS consists of all neurons in the body excluding those in the CNS. Anatomically, the PNS may be thought of as an extension of the CNS in that the functional control centres for the PNS lie in the CNS. The PNS is divided into two systems: the somatic and the autonomic.

- The **somatic system** carries sensory and motor signals to and from the CNS. Thus it innervates sensory mechanisms and striate muscles.
- The **autonomic system** carries impulses that are concerned with the regulation of the body’s internal environment and the maintenance of homeostasis. The autonomic system, therefore, innervates the heart, the smooth muscles and the glands. The autonomic nervous system can be further subdivided into two branches or subsystems (Everly and Rosenfeld, 1981):
  - The sympathetic nervous system is concerned with preparing the body for action. Its effect on the organs it activates is that of generalised arousal.
  - The parasympathetic nervous system is concerned with restorative functions and the relaxation of the body. Its general effects are those of slowing and maintaining basic bodily functions.



CNS – Central nervous system
ANS – Autonomic nervous system
SNS – Sympathetic nervous system
PaNS – Parasympathetic nervous system

Figure 2:9: Structure of the human nervous system

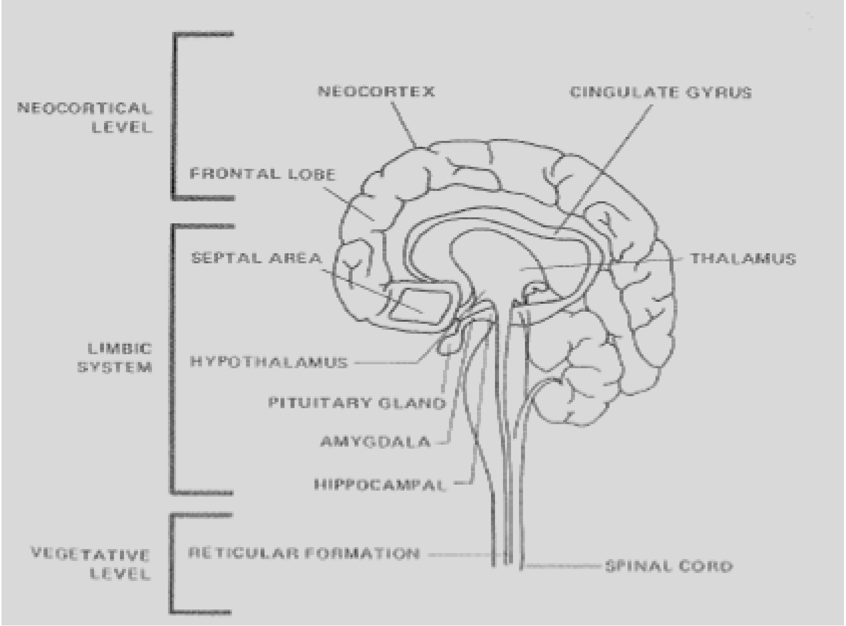


Figure 2.10: The human brain

In this study three functions that are primarily associated with the sympathetic nervous system, i.e. heart rate, heart-rate variability and blood pressure are monitored.

2.8.4.3 The stress response

Everly and Rosenfeld (1981) provide the following description to illustrate the stress response as a reaction to a psychosocial stimulus:

(a) The Initiation of the Stress Response – CNS Mechanisms

If any given, otherwise neutral, stimulus is to evoke a stress response, it must first be received by the sensory receptors of the PNS. Once stimulated, the sensory receptors send impulses along the sensory pathways of the PNS towards the brain. Once in the CNS, collateral nerves from the sensory pathways diverge from the main ascending pathways to the neocortex and innervate the reticular formation. The emotionally-coloured neocortical interpretation is fed back to the limbic system. If the neocortical-limbic system perceives a psychosocial stimulus to be a threat or a challenge, or to be in any way aversive, then emotional arousal will likely result. In most individuals the activation of emotional mechanisms stimulates one or more of the three major psychosomatic stress axes (see Figure 3.6 below). Therefore, we see that stress reactions to psychosocial stimuli result from the cognitive interpretation of the stimuli and the resultant emotional arousal, rather than from the stimuli themselves.

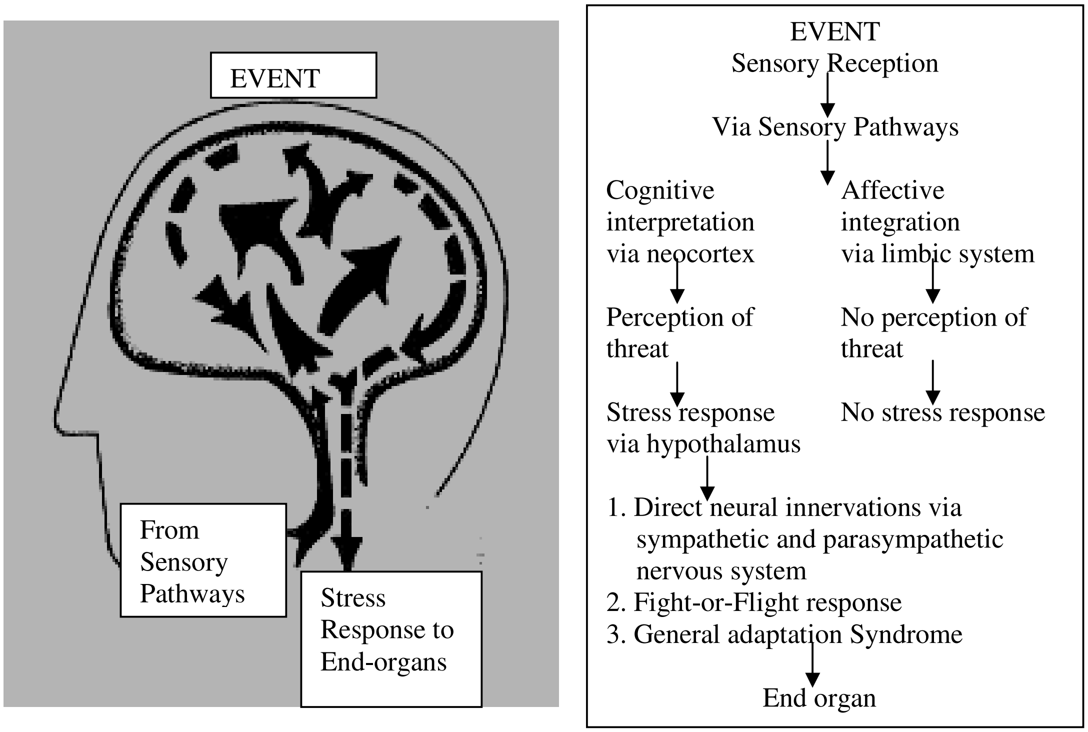


Figure 2.11: The stress response

All stressors will lead to a degree of psychobiological activation. The outcome, however, is strongly influenced by the individual's perception of the stressor. This would consequently influence work performance as well as the degree to which the physiological stress response is activated.

(b) The Psychosomatic Stress Axes

The stress response is of limited duration and is directed towards helping individuals cope with demands. It is therefore adaptational in nature. The non-specific adaptive reactions are directed towards mobilisation of the individual's reserves for energy, modulation of the homeostatic responses, and maintaining a high level of functional activity. However, when demands on individuals become excessive, the neuroendocrine activation associated with the stress response can become chronic and result in adverse effects on performance and health. Performance is, however, often maintained at the cost of health.

The difference between the specific stress response and the non-specific stress response lies in the degree of specificity of the stressor. Stressors that are very specific, such as cold and hunger, have their own specific stress responses that bring the internal homeostasis of the body back to normal through the normal negative feedback mechanisms of the body. However, any stressor that leads to psychobiological activation can give rise to the so-called non-specific stress response that involves, with minor variations, a relative non-specific neuroendocrine activation. This so-called non-specific stress condition or response can be seen as a new homeostasis meant to enable the individual to cope, mentally and physically, with the demands that initiated the response. Although the non-specific stress response involves virtually the whole body, as previously mentioned, two major neuroendocrine systems are involved in the peripheral expression of the response, i.e., the sympathoadrenomedullary (SAM) axis and the hypothalamo-pituitary-adrenocortical (HPA) axis (Viljoen, Claassen and Hazelhurst, 2004. See Appendix D).

The SAM and HPA axes are the two main stress axes that are activated after prolonged exposure to a stressor(s) and have therefore formed the basis of the physiological measurements of the verification study. The non-specific stress response varies widely from individual to individual (interpersonal differences) and even within one individual at different times (intrapersonal). This makes it very difficult to be definitive about the stress levels and stress responsivity of any individual. The discussion that follows will be limited to those physiological parameters that were possible to measure within the limitations set by the brief of this study, i.e., a non-intervention study within the daytime working hours of the TCOs.

(c) The Neuroendocrine or Sympatho-adrenomedullary (SAM) Axis

The SAM axis consists of the sympathetic nervous system and the adrenal medulla.

When mass activation of the SAM axis occurs it is referred to as the fight-or-flight response. The effect is to aid the individual in coping with environmental and other demands and entails a mobilisation of the body to prepare for muscular activity in response to a perceived threat. This mechanism allows the organism to either fight or flee from the perceived threat.

The pivotal system in this response is the adrenal medulla, which can be seen as an extension of the sympathetic nervous system. Stimulation of the adrenal medulla releases adrenalin and noradrenalin (catecholamines) into the systemic circulation. The effect of this release helps to increase energy availability and, in addition, leads to an increase in cardiac output, an increase in blood pressure, and a differential increase in blood flow (Viljoen et al., 2004. See Appendix D).

It is important to note that the secretory response of the adrenal medulla to stressors is dependent on the previous stress history of the individual. For instance, with exposure to the same stressor each day over an extended period of time, subsequent exposure to the same stressor may induce a lower adrenal medullary response. This response could be interpreted as habituation. However, if the same individual who has been stressed by a particular stressor over an extended period of time is suddenly, in addition to the previous stressor, exposed to another, novel, stressor, the adrenal medullary secretory response may be significantly larger. This response is known as sensitisation (Viljoen et al., 2004. See Appendix D).

In terms of TCOs and their stress levels or, more specifically, their adrenal medullary stress responses, it would imply that they may adjust to the same workload when exposed to it over a period of time (habituation), but that the introduction of significant changes could lead to higher stress levels than before (sensitisation).

According to Everly and Rosenfeld (1981), specific somatic effects that have been observed in human beings as result of activation of the SAM axis in response to psychosocial stimuli are:

- increased arterial blood pressure
- increased cardiac output
- vasoconstriction
- increased muscle tension
- increased cholesterol levels.

(d) The Endocrine or Hypothalamo-pituitary-adrenocortical (HPA) Axis

The most chronic and prolonged somatic responses to stress result from the stimulation of the endocrine axes. Activation of the HPA axis during stress leads to an increase in, along with other neurotransmitters, cortisol, which in turn leads to increased access to energy stores as a result of conversion of other substances like lipids to glucose (Viljoen et al., 2004. See Appendix D).

The circadian rhythmic secretion of cortisol is characterised by peak secretion in the early morning just before getting up and during early morning activity. It levels off later in the day but may increase as a result of meals or any of a variety of stressors. One of the major effects of stress, including heavy workloads, unpleasant working environments or even continuous shift changes between day and night shifts, is, in fact, an alteration of the circadian cortisol secretory pattern.

Stress has several effects on the circadian rhythmicity of cortisol. Acute stress can cause acute increases in cortisol levels, which usually return to normal again within two hours. This is especially marked in the case of aversive stressors. Similarly to the SAM axis, incidents of acute heterotypic stressor application during periods of chronic stress alter the response. It would seem that acute stress superimposed on chronic stress is dependent on the familiarity with the type of stressor. It appears that the response to an acute stressor in the chronically stressed individual is less than expected if the stressor is homotypic (a repetition of what caused the chronic stress response). In contrast, if a heterotypic stressor were applied to a chronically stressed individual, the response may be more intense. It is however possible that cross-tolerance to stressors may develop, especially with stressors that involve the same neurological pathways.

In theory this can be extrapolated to the working situation where high cortisol levels would not be significantly increased by increased levels of homotypic stressors such as increases in the amount of work to which the individual is accustomed. In contrast, the response can be exacerbated when atypical stressors are encountered. The implication is that individuals with chronically high cortisol levels will not show the expected increase from basal trough values when stress levels increase. This could be ascribed to the fact that the trough values are chronically higher than normal (Viljoen et al., 2004. See Appendix D).

The effects of chronically high cortisol levels will again be referred to in the discussion of the verification study results.

A discussion of the human stress response and the accompanying measures of stress associated with high workload would be incomplete without reference to allostatic load. In Addendum D the potential effects of allostatic load in the context of TCOs' job demands, and the resulting high mental workload, are discussed.

(e) Allostatic Load

Various types of allostatic responses to stress can occur. With the normal acute allostatic response, which is aimed at being of benefit to the individual, the response that usually occurs includes activation of the SAM axis and the HPA axis with subsequent increases in heart rate, blood pressure, blood sugar and other physiological changes. The response is terminated after the stressful event, and the physiology of the body returns to normal. As long as the allostatic response is limited to the period of work, or the demands on the individual are not excessive, protection via adaptation predominates over any adverse consequences. However, exposure to the allostatic response over extended periods of time can have psychopathological consequences. Chronically stressful situations and other factors may lead to abnormal allostatic responses which can, in turn, give rise to a high allostatic load.

The allostatic load can be seen as similar to the effect of the long-term or chronic so-called non-specific stress response. It is the price paid for the repeated or chronic activation of the allostatic response – a response which in the short term is intended to be advantageous. Another way of putting it would be to define it as the wear and tear that results from chronic over activity or under activity of the allostatic systems. The allostatic load is the product of accumulated stressors over an extended period. This includes the effects of minor day-to-day stressors and more dramatic or traumatic events. There can be no doubt that type of work, the working environment and working hours can be seen as major determinants of stress. In this the perceptions of individuals play a significant role. It is one of the major factors that determine whether they will become sensitised or habituated to a certain stressor, i.e., how they will cope with the different day-to-day stressors and for how long their neuroendocrine systems will remain activated. Other factors that can be seen as instrumental in modulating the characteristics of the stress or allostatic response are the general physical health and lifestyle of the individuals, i.e. exercise, diet, alcohol intake and smoking (Viljoen et al., 2004. See Appendix D).

From a medical point of view, as well as in terms of common human interest, an important aspect of mental load would be to establish whether it is possible to identify loads that may lead to pathological consequences.

The relationship between health and stress is both complex and confusing. When coping is possible, or when individuals evaluate the situation as one they can cope with, there is no evidence of pathology or of bodily changes that may be assumed to lead to pathology. When coping or control appears impossible, pathology develops, at least in acute animal experiments (Ursin and Ursin, 1979).

Therefore, if it is impossible to measure mental load per se, psychosomatic complications could indicate prolonged exposure to overload, even if it is indicated simply by the failure of coping processes. These psychosomatic complications depend not only on the mental load but on competence and subjective evaluations of load and capacity.

From a physiological viewpoint, a mental load must be assumed to be a load on processes within the central nervous system (CNS). Physiological indicators pick up activation rather than information processing. Activation is the end result of a wide range of psychological processes (Ursin and Ursin, 1979).

The use of physiological indices therefore rests on the assumption that it is possible to assess the amount of effort expended by operators in meeting the demands of the task. Changes in activation levels of subjects will ensue and these can be measured. The literature uses concepts such as arousal, activation, and effort interchangeably in this regard.

Activation theory is still a powerful model for the explanation of physiological and psychological mechanisms in wakefulness, sleep, and emotional states. This theory states that when there is information transmitted through the usual sensory pathways, there are also impulses sent directly to the reticular formation of the brainstem. The activation is not restricted to cortical activation, but also involves activity in the autonomic nervous system, the somatomotor system, and endocrine system. With improved methods for determining the plasma levels of hormones, it has become increasingly clear that the whole or at least a very large part of the endocrine system is subject to the influence of psychological factors. These phenomena are easily treated within activation theory (Mulder, 1979).

It is suggested that whenever physiological indicators are used to measure mental load, psychological and social factors must be assumed to be of decisive importance for the values obtained. Physiological indicators, therefore, pick up what may be called ‘emotional’ factors, or simply ‘activation’. Activation depends on whether individuals expect to cope with their situation or not. The methods do not measure the information load. They give information on how individuals estimate the load, in particular whether they think they are able to cope with it, and to what extent failure or success is important to them. Therefore it is not the load that is measured but their evaluation of the load and their evaluation of their capacity to cope with that load. This depends on experience as well as personality characteristics.

The assumption that underlies the use of psycho-physiological measures of stress and mental workload is simple: as workload increases, so a corresponding increase in operators’ levels of arousal is reflected in the activity of the autonomic nervous system. This level of arousal can be recorded by a number of psycho-physiological techniques. While no general pattern of changes in the various psycho-physiological variables in response to known changes in workload has emerged, it has been shown that certain measures demonstrate relative specificity to different workload components (Megaw, 2005).

Physiological measures such as heart rate, blood pressure, respiratory rate, eye blink, and pupil diameter are related to autonomic nervous system responses and hence are not under voluntary control or subject to deliberate bias or subjectivity. These measures can be taken without intruding on the primary task and can be indicators of mental workload and the stress emanating from the workload. They could also be especially helpful in the process of validating other techniques (Kramer, 1991).

Some of the physiological measures of mental workload that have yielded positive research results in mental workload assessment are described briefly below:

2.8.4.4 Measures of Brain Function

(a) Event-related Potentials

Event-related potentials (ERPs) are fluctuations in the endogenous activity of the nervous system that is recorded in response to environmental stimulation and that are associated with psychological processes or in preparation for motor activity. Various ERP components have been taken as indicative of information processing activity and changes in mental workload (Hancock, Meshkati, Robertson and Robertson, 1985). The ERP is time-locked to the stimulus event. Workload inferences are based upon the amplitude and latency elements of the elicited ERP wave (Meshkati et al., 1995).

ERPs represent the most valid physiological measure of workload, but limitations such as the considerable supporting instrumentation, sensitivity to environmental electrical noise and to the movement of the individual, the requirement of trained personnel for operation and interpretation, and sensitivity to individual differences, which require calibrations for each individual operator, make it a difficult measurement in practical terms and intrusive on the work environment.

(b) Electro-encephalograph Changes

Experiments have shown that mental activity affects the frequencies of the electro-encephalograph (EEG) but the relevance of this is confused by large inter- and intraindividual differences (Meister, 1985). This approach is extremely intrusive and creates an artificial work environment.

2.8.4.5 Measures of Heart Function

(a) Blood Pressure

Blood pressure is the outward force exerted against the walls of blood vessels. Four pressures were considered for the purposes of this study, i.e., systolic pressure, diastolic pressure, pulse pressure and mean arterial pressure. Increased blood pressure is a major complication of a stressful lifestyle. It is therefore a good reflection of the accumulated wear and tear of stress (allostatic load) on individuals (Viljoen et al., 2004. See Appendix D). However, the measurement of blood pressure as an indicator of acute stress often does not yield the information required to assess the magnitude of the stress response (Kaplan, 2000).

(b) Heart Rate

Heart rate is largely the product of direct or indirect sympathetic nervous system responses and the involvement of neurohormones. In contrast to blood pressure, the sensitivity of heart rate to stress-induced central nervous system activation is generally considered to render a justified index of acute psychological stress (Baker, Suchday, and Kranz, 2000).

Many examinations of heart rate fail to consider heart rate within the context of cardiovascular control mechanisms. When the environment (temperature, acceleration related to flight operations) is not benign, it may be impossible to discriminate cognitive effects on heart rate from environmental effects associated with physical effort and thermoregulation (Gawron et al., 1989).

Generally, increased heart rates are associated with increased levels of workload. Such increases have been found to vary with workload in several studies with pilots. However, not all investigators report consistent findings of heart-rate changes that occur with differences in workload (Wierwille et al., 1985). This inconsistency of results has caused a number of investigators to abandon simple heart-rate measures and to look instead at the variability of the heart rate as a possible measure of cognitive workload (Wilson and O'Donnell, 1988).

(c) Heart-rate Variability

Measures pertaining to heart rate and their derivatives are currently the most practical physiological method of assessing mental workload. Among such measures perhaps none is more thoroughly investigated than that of heart-rate variability (HRV). Several empirical investigations attest to the strength of this assertion (Steptoe, 1981 and Meshkati, 1988).

Heart-rate variability (HRV), which is based on small changes in the beat-to-beat or interbeat intervals (also referred to as R-R intervals), provides a non-invasive indication of autonomic function. HRV is normally measured by variations in the interbeat interval variability in heart rate. It is generally thought that increased sympathetic stimulation decreases heart-rate variability and that increased parasympathetic stimulation increases heart-rate variability (Woo, Stevenson, Moser, Trelease and Harper, 1992).

In a study on air traffic controllers, whose tasks show a number of similarities with those of TCOs, Kalsbeek (1971), using interbeat intervals, noted a gradual suppression of the heart-rate irregularity resulting from increases in the difficulty of the task.

Wierwille (1979), whose research also involved aircrew mental workload, is supportive of the above argument. He concluded that heart-rate variability decreased with increased mental load.

Several studies have since been conducted on the use of HRV as a quantitative index for mental workload (Kumashiro, 2005). The variance of the inter-beat interval times with increasing taskload has been confirmed in a number of experiments (Mulder and Mulder, 1987). Veltman and Gaillard (1998) investigated the sensitivity of physiological measures to mental workload on pilots in a flight simulator. They found that HRV was sensitive to large changes in task difficulty. Consequently, it was posited that such a measure could be used to reflect mental workload.

The following conclusions can be drawn from a review of the literature:

- Heart rate generally does not change with mental load. In some instances stress might cause a change in heart rate.
- Heart-rate variability may decrease with an increase in mental load.
- Blood pressure may increase with stress but is not a reliable measure of workload.
- The use of spectrum analysis of heartbeat intervals shows promise as a means of estimating workload.

(d) Respiratory Rate

This is one of the easiest measures to record and has been used extensively in studies involving emotion, stress, arousal and mental load (Meister, 1985).

2.8.4.6 Eye Function Parameters

(a) Eye Blink Rate

Eye blink duration and eye blink frequency are good indications of the onset of fatigue and/or when a person is under stress. It has been found that eye blink rate decreases under conditions of high workload both in driving tasks and in flight tasks (Hankins and Wilson, 1998). Laboratory studies have demonstrated that tasks requiring attention, especially visual attention, are associated with fewer blinks and blinks of a shorter duration (Wilson and O’Donnell, 1988). Infrared eye blink detection technology makes this a practical measure to use because it is not intrusive on the performance of workers. Developments include systems that detect and measure both eye blinks and slow closures since both types of eye closure reveal information about operators' alertness levels (Knipling and Wierwille, 1994).

(b) Pupil Diameter

Several investigators have observed that the diameter of the pupil correlates quite closely with the resource demands of a large number of diverse cognitive activities (e.g., mental arithmetic, short-term memory load, air traffic control monitoring load, and logical problem solving). This diversity of responsiveness suggests that pupilometric measures may be highly sensitive, although as a result, it is undiagnostic. Pupil diameter reflects demands imposed anywhere within the system. Another disadvantage is that relevant pupil changes are in the order of tenths of a millimetre, which means that accurate measurement requires considerable constraint of the head and precise measuring equipment. In addition to this, changes in ambient illumination must be monitored since these also affect pupil diameter. Because of its association with the autonomic nervous system, the measure will also be susceptible to variations in emotional arousal (Wickens and Hollands, 2000).

(c) Visual Scanning

Visual scanning, the direction of pupil gaze, although used as a measure of attention allocation, can contribute extensively to workload measurement in two different ways. Firstly, dwell time can serve as an index of resources required to extract information from a single source. Secondly, scanning can be a diagnostic index of the source of workload within a multi-element display environment (Wickens and Hollands, 2000).

2.8.4.7 Body Fluid Analysis

(a) Cortisol

A novel physiological approach to assessing mental workload was the investigation into psychoneuroendocrine responses to mental workload by Hyyppa, Aunola, Lahtela, Lahti and Marniemi (1983). They were able to find a significant decline in the cortisol and prolactin levels of subjects undergoing psychologically demanding achievement-orientated tasks.

Stress hormones measured in the urine, blood or saliva can be a good control measure to determine if operators have experienced stress over a particular period of time. This technique is reliable but costly, and often employees object to blood and urine samples being taken for analysis for fear that other substances may also be detected. Recently however saliva test kits have been made available to measure stress hormones in saliva samples.

Benefits and Shortcomings of Physiological Measurements:

Physiological indexes have two distinct advantages -
(1) They provide a relatively continuous record of data over time.
(2) They are not intrusive on primary-task performance.

In terms of their disadvantages, physiological measurements often require the attachment of electrodes (for ECG measurement) or some degree of physical constraint (pupilometric measurement), and therefore they are not unobtrusive in the physical sense. These constraints may influence user acceptance. Many physiological measures have a further potential disadvantage in that they are, generally, one conceptual step removed from the inference that the system designers would like to make, that is, workload differences measured by physiological means must be used to infer that performance breakdowns will result or to infer how operators will feel about a task. Secondary- and primary-task measures assess the former directly, whereas subjective measures assess the latter (Wickens and Hollands, 2000).

When using physiological measures, care must be exercised in their selection and application. In general, the sensitivity of these measures is task dependent. Physiological measures might be more useful in assessing the effects of time on task and task strain. Also, in terms of intrusiveness, physiological measures have been fairly successful indicators without influencing the primary-task performance of operators (Meshkati et al., 1995).

The use of physiological measures as indicators of mental workload is influenced by a combination of several factors, such as the cost of both the hardware and software required to operate the equipment, the training level of the personnel who administer the physiological tests, environmental conditions at the workplace and the willingness of employees to be connected to a physiological recording mechanism. Due to problems caused by various combinations of these factors at the present time, most physiological measures are impractical for use in the assessment of mental workload in complex machine-person systems. However, technological innovations, such as miniaturisation, telemetry and data-logging have reduced some of the problems, and as a result, the use of some of the techniques has become more feasible.

In this section the different groups of mental workload measurement techniques have been discussed. The purpose of this review was to obtain an understanding of the existing techniques and to consider which of these techniques would be appropriate and practicable for use in the Spoornet train control context.

From the available research it is clear that the various mental workload measurement tools have been applied in a variety of research environments. From the literature reviewed it is evident that the period between the late 1970s and late 1980s was the most productive in terms of mental workload research. During this period most of the papers and research results published were based on commissioned and sponsored research in military, aviation and aerospace contexts. Many reports, especially on the implementation of specific models or tools, are classified and not available in the public domain. A further significant observation is that since 1990 very little research has been conducted. The only papers available are on improved physiological measurements, and no new research has been conducted or reported on that focuses on new ways of assessing mental workload.

The only previous South African study on mental workload, undertaken by Cilliers (1992), was a doctoral study performed in the South African Air Force. The purpose of the study was to select a number of tools to assess the mental workload of fixed-wing fighter pilots. No new tools were developed and the study essentially entailed the correlation of various known workload assessment tools (all of which have been discussed in this chapter) with one another. No evidence of the implementation of the selected tools could be found, unless this is contained in a classified report that is not currently available in the public domain.

Another shortcoming is that the existing literature does not provide sufficient guidelines for the application of the tools in an environment other than for that which the specific tool was developed. Furthermore, there is limited information on the nature of tasks and simulation environments in which the research was conducted. These factors limit the degree to which the techniques can be applied in an environment unrelated to the research environment.

The groups of mental workload measurement tools and techniques reviewed here each have reported advantages and disadvantages. In conclusion, each of these groups will be critically evaluated against the set criteria and in terms of their appropriateness and applicability in the current context.

The following criteria for a mental workload assessment tool were identified as important in the Spoornet context:
Sensitivity
Selectivity/Validity
Intrusiveness
Operator acceptance
Affordability
Implementation requirements

The reasoning behind not utilising a simulated environment for the development and testing of a new mental workload methodology needs to be explained. The literature has many reports on the use of simulated environments for testing new methodologies, and it was the obvious choice for this study. Upon inspecting the training facility for the radio train order environment, it became evident that real-life scenarios could not be simulated and that fidelity would be a major concern. RTO training is the only training in the train control environment that cannot be simulated, unlike the situation where TCOs work with colour light signals that are monitored on a panel. To overcome this problem from a training perspective, newly qualified TCOs are never sent to an RTO section until they have gained sufficient experience in other environments. They then receive supervised ‘on the job’ training in an RTO control centre and are certified *in vivo* while under supervision before they can work in the RTO environment. The use of a simulated environment has therefore never been an option for this study. This has resulted in a number of limitations on methods that could have been used but due to the safety criticality of the tasks of TCOs had to be abandoned.

The rationale for not using existing mental workload measurement methods is provided below.

2.8.4.8 Measures of Primary-task Performance

Considering the context, this group of measurements would violate the criterion of intrusiveness. Given the safety responsibilities of TCOs, this is a critical requirement. The criterion of sensitivity would also not be satisfied as the task demands of TCOs are not consistent but rather occur in surges. For this reason degradation of performance would be difficult to detect.

2.8.4.9 Secondary Measures of Spare Capacity

This group of measurements, due to its nature, poses a safety risk and, in the context of the TCOs’ tasks, would be intrusive. This group was the first to be rejected by Spoornet management and trade unions when the option of simulation was ruled out.

2.8.4.10 Subjective Rating Techniques

The very nature of subjective ratings, i.e., assessing one's own feelings of fatigue, mental effort, anxiety, stress load, or frustration, disqualified the use of this group in the Spoornet context. The reason for this decision was twofold:

- Spoornet is a highly unionised work environment. Union members will only participate in a study of this nature upon receiving a mandate from their union leaders. All trade unions were therefore involved as stakeholders from the inception of the project. Spoornet has been subject to retrenchments since the early 1990s, and during the period when this project was executed (2001-2004) the process had not been finalised and retrenchments were ongoing. Union leaders were adamant that their members should not be compromised in any way. Completing a checklist or rating scale on how employees experienced their shift, constituted such a compromise to union leaders. They were concerned that employees might be judged as incompetent and could therefore be retrenched on the basis of their experienced workload as captured on a self report form. Despite guarantees from the researchers that employees would not be identified, this group of measurements was not acceptable and therefore did not satisfy the user acceptance criterion.
- Secondly, subjective rating techniques are essentially the rater's perception of the difficulty of a task. Spoornet management was not prepared to accept the perceptions of individuals as a basis on which to make operational safety decisions, especially if such perceptions were tainted by concerns regarding job security. The requirement at the outset was that an objective assessment of mental workload was required. Furthermore, it should be quantified in order to prevent individual interpretation. Simply put, a number – and preferably a cut-off number within the limits of which mental workload is considered safe – should be attached to the mental workload of a specific TCO within a specific train control centre.

As indicated earlier in this chapter, the NASA-TLX is considered an industry standard as far as subjective ratings of mental workload is concerned. There is however no such standard for objective ratings. In the planning phase of the project it was decided to validate the newly developed MWLI against physiological measurements, as well as against the NASA-TLX, because of its status as an industry standard. In the verification phase of the project, which posed a tremendous number of challenges in terms of logistical arrangements (which is referred to in a later chapter) it became clear that the NASA-TLX could not be used. The reason for this was the fact that various train control centres had different shift patterns, depending on the size and activities at the centre. To simplify the measurement process, it was decided that all measurements would be taken between 06:00 and 14:00. For all centres this meant the start of a new shift (06:00) but not necessarily the end of a shift as some centres had 8-hour shifts and others 12-hour shifts. In order for the NASA-TLX to not be intrusive and pose a safety risk in this context, it could only be administered at the end of a shift. For many of the centres, the end of a shift meant 18:00 and the measurements were only taken over an 8-hour period starting at 06:00, irrespective of whether the shift was 12 hours in total. The reason for this decision was to ensure that unrelated factors, such as circadian rhythm, which would affect cortisol measurements, would be eliminated. Physiological measurements, as proven objective indicators of whether a workload problem exists, provide a more valuable validation methodology in this context and the NASA-TLX was therefore eliminated from the verification phase.

From this overview of the existing body of research in the field of mental workload it is evident that there is clarity as far as the factors that contribute to mental workload are concerned. A practicable index for the measurement and prediction of mental workload is, however, not available.

With this challenge at hand, and taking into account the criticisms lodged against some of the existing methods, an objective, user-friendly, non-invasive method of measuring the mental workload of train control officers at Spoornet had to be developed that could be used on operators while they were performing their tasks. The method had to be reliable and valid as well as verifiable and had to comply with the other criteria referred to previously for the selection of mental workload techniques.

The existing mental workload measurement techniques have to be considered in the context of the study in order to understand their applicability. This context is consequently provided.

2.9 CONTEXTUALISATION OF THE STUDY

2.9.1 Existing research

The problems that were faced in this study related to the existing body of knowledge as well as to the context within which the research was undertaken.

Since the inception of mental workload research in the 1970s, researchers have used a variety of definitions of mental workload. There is however not one single definition of mental workload that could be used to guide ongoing research. The definitions available have also evolved over the years, and engineering, psychological and physiological experts have all provided paradigm-specific perspectives.

Meshkati (1988) (Figure 2.5) has developed a single model of mental workload that incorporates all the factors relating to workload. It incorporates most of the definitions used by researchers and highlights the development processes involved. Although this model dates back to 1988, there is no more recent integrated model and it was therefore considered a useful basis for this study. As mentioned earlier, the literature indicates that no single definition of workload will suffice due to the complex nature of the concept. This model captures this multi-definition approach and also emphasises the multidisciplinary nature of mental workload. This study does not accept any single definition of workload as the paradigm from which a mental workload assessment model should be developed. The model considers the task variables related to workload, individual moderating factors, and various measurement techniques, although no specific recommendations are made for the objective assessment of mental workload.

The limitations in the existing research, which have also necessitated this study, relate to the fact that there is no clarity on measurement techniques that fulfil the set criteria. There is no readily available technique that could be used across a variety of contexts. The research also indicates that techniques applied and reported on in various contexts are situation and context specific.

In order to find a solution for mental overload in the train control environment, the air traffic control (ATC) domain seemed a logical benchmark.

Nunes (2003), in a study on assessing the impact of predictive aids on controller performance, came to the conclusion that the traditional need to design interfaces in ATC based on direct visualisation have been grounded in the notion of reducing controllers' cognitive demand. His study confirmed previously expressed concerns that such aids can have detrimental effects on the underlying mental model, which in turn can compromise operators' problem-solving and learning abilities. In this particular study the NASA-TLX was used as a means to measure mental workload experienced.

From this study it seems that the reduction of mental workload through the provision of predictive aids might compromise the performance of controllers. Personal communications with the South African Air Traffic and Navigation Services (AT&NS) (2004) indicated that no specific measurement or tool is used to detect and manage mental overload. Rather, mental overload is managed through limited exposure to a high load by keeping shift lengths to four hours over peak periods.

A study by Averty, Collet, Dittmar, Athenes and Vernet-Maury (2004) makes an interesting contribution to ATC mental workload measurement. This study reports on the development of the traffic load index (TLI). Mental workload assessment studies of ATC have used either objective measures, i.e., numbers and distribution of aircraft, or subjective factors, such as self-imposed performance and stress levels, with mixed results. The number of aircraft (N) was the most common index used. Due to safety constraints, subjective aspects of workload (strain) have never been assessed in real time during actual work sessions. In the study quoted here, it is hypothesised that N is not a perfect index and that a better index of mental load measurement could be obtained if the interventions carried out by controllers were integrated in the index. In the study data were collected on communications with pilots, radar material and the number of aircraft (a distinction was made between aircraft monitored and aircraft with a control problem). Time pressure and uncertainty was also factored into the TLI. This data was then quantified. Controllers were also requested to complete the NASA-TLX during the experiment. The experiment took place in the field but arrangements were made to complete the NASA-TLX between sessions. The results of the TLI were compared with N and the NASA-TLX. TLI is shown to be more highly correlated to TLX than N to TLX. Conversely, the correlation between TLX and N was lower, meaning that N does not exactly reflect the perceived workload. Future work with the TLI entails demonstrating its validity. Correlation with psycho-physiological variables to validate TLI is envisaged.

The similarities between the study by Averty et al, (2004) and the study reported on in this thesis are noticeable. These similarities are summarised as follows:

Table 2.5: Similarities between TLI and MWLI development

Factor	**TLI development**	**MWLI development**
Existing measurement(s) of mental workload (crude)	Number of aircraft (N)	Spoornet formula
Safety risks relating to field work – limits study	Air traffic control	Train control
Methodology:		
1. Recorded communications quantified	1. Communications with pilots	1. Communications with train drivers and others
2. Operational records	2. Radar information	2. Train schedules and plans
3. Objective data: Actions quantified	3. Number of aircraft	3. Number of authorisations and data transactions
4. Subjective/additional data	4. Emotional processes linked to conflicts and time pressure	4. Moderating factors
5. Added measurement of workload	5. NASA-TLX	5. TLA
Validation	Psychophysiological measurements envisaged	Psychophysiological measurements used on limited population

Cilliers' (1992) work refers to two other South African studies, both in the military aviation context. The reports are referenced as confidential reports and were never published in the public domain. Cilliers' study was also conducted in the field of military aviation and she therefore had access to these confidential military reports. Besides these two confidential reports, Cilliers' (1992) study is the only study on mental workload that was ever undertaken in South Africa. Cilliers quotes and comments extensively on the existing body of knowledge and come to the same conclusion, namely, that no specific techniques are recommended, no guidelines exist on the tasks that were included in research projects, and few guidelines exist for simulated environments.

2.9.2 Research in the Rail Environment

Mental workload research in the rail environment has never been undertaken in South Africa. Internationally, a few published studies are available. The work undertaken by Devoe (1974) is one of the most comprehensive and informative studies undertaken for the purpose of identifying and understanding the tasks of railroad train dispatchers (the equivalent term for train controllers) and the environment and context in which they occur.

The past few years have seen a resurgence of interest in rail human factors and specifically mental workload. In the USA the Transportation Research Board (TRB) is the main contributor to research in this field. The author of this paper is a member of the TRB Taskforce on Railroad Operational Safety and therefore has access to contract research reports.

As mentioned before, between the late 1980s and 2003 very little new work on human mental workload was published. Previously developed subjective measurements, especially the NASA-TLX, have become industry standards, specifically in the aviation industry.

During 2001 two studies relating to dispatcher workload were sponsored by the Federal Railroad Administration (FRA) and comprehensive reports were produced (Popkin, Gertler, and Reinach, 2001 and Reinach, 2001). Elements of both these reports were incorporated in developing the MWLI.

The purpose of the Popkin, et al. (2001) study was, in the interest of railroad operational safety, to create a better understanding of the dispatching environment and its associated levels of workload, occupational stress and fatigue. The project goals were:

- Identify the sources and magnitude of workload, stress and fatigue associated with the railroad dispatcher's job and working life
- Determine any related health or performance effects

- Refine procedures for measuring workload, stress and fatigue in the dispatcher’s workplace.

There were three sources of workload data, including an observational technique, based on the Task Analysis Workload method (TAWL), subjective ratings and activity-count data. Salivary cortisol was used as a physiological measure of stress while actigraphy was used to record sleep patterns. The modified Task Analysis Workload (mTAWL) measurements provide a means to gauge variation in workload over the course of the shift.

The key findings of the Popkin et al. (2001) study were:

- Subjective workload was moderately associated with the reported number of trains dispatched, regardless of shift or location. Perceived stress also related to the number of trains dispatched.
- The mTAWL is labour intensive and not suitable for a research study.
- There was little evidence of high levels of stress from either subjective stress ratings or from salivary cortisol levels. These levels were well within normal levels for adults. These results should not be interpreted as an indication that workplace stress does not exist. It is more likely that data was collected too infrequently in this rapidly changing environment and may not have captured the changing workload and related stress. (The authors of the report noted that given that the data was collected from a small, non-randomly selected sample of dispatchers (N=20) from two dispatching centres, the results should be carefully interpreted.)

A technical memorandum (Reinach, 2001) was produced in response to concerns raised by a number of FRA audits over the safety of the U.S. rail network. Specifically, concerns were raised about the absence of a method that could be used to reliably collect dispatcher-related data (e.g., workload) as part of FRA safety and compliance inspections and audits of dispatching centres. A major consideration in collecting taskload data, which is the basis of an assessment tool, is that the tasks must be observable, quantifiable, quick, easy and unobtrusive to collect. The ultimate goal of the research was to produce a tool that FRA field personnel and others could use to reliably and quickly obtain taskload data. The technical memorandum was the first step in the development of a dispatcher taskload assessment tool. The rationale and ultimate goal of this particular study shows significant similitude with the Spoornet study.

Reinach (2001) identifies 67 unique dispatcher tasks, which were grouped into six categories. A number of factors that were expected to affect dispatchers' taskload were also identified. A number of recommendations for further activities were made before a taskload assessment tool could be developed.

Reinach (2006) elaborated on the above study and developed a performance model of railroad dispatching. It is evident that subjective measurements developed and refined for the aviation industry could not be superimposed on the rail industry. Taskload in this context was defined as the average time demanded of a dispatcher in carrying out all job-related tasks at a dispatching desk over one shift. The methodology of the study involved the collecting of taskload data through observations and questionnaires. The same 67 tasks from the 2001 study, identified and organised into six top-level task categories, were used. These were:

- actuation of signals, switches, blocking devices and bridge controls via centralised traffic control (CTC) or computer-aided dispatching systems;
- issuance and cancellation of dispatcher-authorised mandatory directives;
- granting of other track-related permissions, protections and clearances (non-mandatory directives);
- performing non-movement authority or non-permission/protection/clearance communications (This generally involves advisories, coordinating activities and the exchange of work-related information.);
- Performance of general record-keeping tasks;
- Review of reference materials.

These tasks represent the gamut of possible railroad dispatching activities across the United States, regardless of the dispatching technologies used or the nature or size of the operation. A number of other factors that either affect dispatchers' taskloads, or can be used to describe the circumstances in which taskload is measured, were also identified. Some factors are internal to the dispatchers, while other factors are external. In the

Spoornet study, similar factors were identified and referred to as *moderating factors* (see Chapter 3).

The following factors were identified in the TRB study:

- track-related factors
- railroad operation-related factors
- dispatcher-related factors
- other factors

The goal of this TRB study was to identify dispatcher tasks and data collection methods that would support the development of a dispatcher *taskload* assessment tool that could reliably measure observable dispatcher activity at different dispatching desks. This goal is very similar to the goal of the Spoornet study. Reinach's study concludes that:

- Taskload assessment has its limitations, particularly due to the highly cognitive nature of the job. The physical activities of dispatchers are observable but not cognitive activities.
- It is possible that dispatchers' performance may not be easily measured using only the physical work (taskload).
- The data gathered as part of this research provided a better understanding of the job of railroad dispatchers and contributed to the development of a preliminary model of dispatcher performance and safety that incorporates both physical and cognitive aspects of the job (see Figure 2.12).
- The proposed model, when fully developed and validated could provide a broader understanding of dispatchers' jobs. The model could be used to support a number of activities, including technological developments that might facilitate safety, more efficient dispatching, identification and development of criteria for more effective training programs, and monitoring the effects of changes in technology on the task of dispatching.

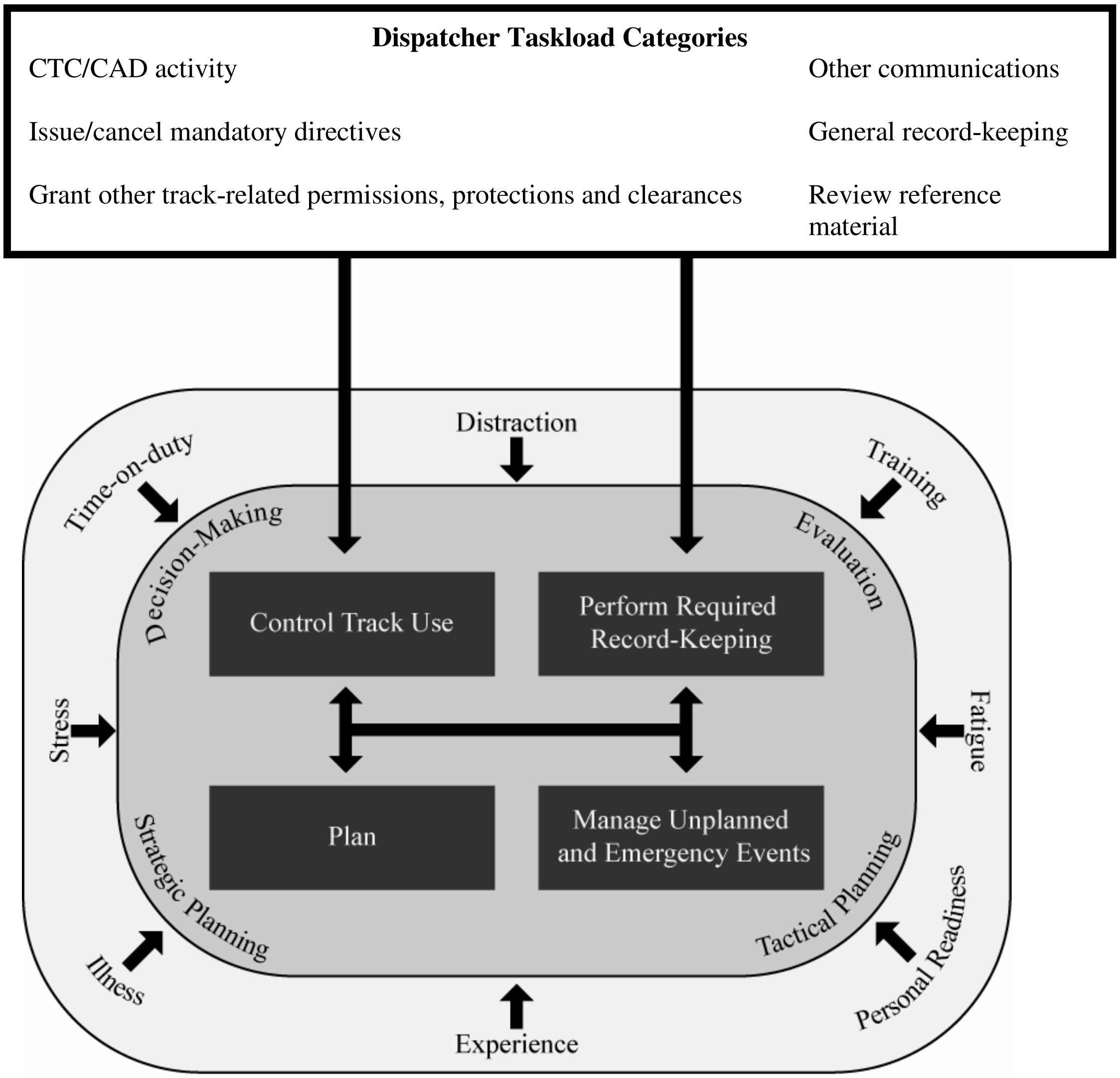


Figure 2.12: Preliminary Model of Railroad dispatching (Reinach, 2006)

In Europe the main contributor to human factors research in the rail environment is the Centre for Rail Human Factors at the University of Nottingham, (UK), which supports the Rail Safety and Standards Board in Britain with projects and research in the field of human factors. Two studies on mental workload from this centre were published in late 2005.

Pickup, Wilson, Sharples, Norris, Clarke and Young (2005) reiterate the same concern as the FRA, namely that mental workload in the signalling (or train control) environment is an area that needs to be better understood and practically assessed. They propose a conceptual framework and a suite of workload tools.

Pickup et al. (2005) acknowledge that mental workload is a complex field of study that starts with various definitions and a plethora of associated measurement methods. A further difficulty is that researchers have often employed one of a number of well known (self-report or subjective) MWL measurement methods, in the hope of setting their results against some norms, but as a result may be trying to use a benchmark with a measure that has been created for an entirely different purpose in entirely different circumstances.

The context of the work of Pickup et al. (2005) is very similar to the Spoornet context. A brief account of their research findings is given here in order to contextualise the Spoornet study.

The motivation for the research was the need of Network Rail (the owners of the rail infrastructure in the UK) to understand and measure the workload of signallers (elsewhere called dispatchers or controllers) in the context of technical and organisational system change. Network Rail wanted to derive a 'function complexity index' based on the principle that there should be a (ideally linear) relationship between the number of inputs received and the activities of signallers and their subjective perceptions of mental workload. This early approach was, thus, based on examination of the relationship between the demand for resources imposed by a task and the ability to supply these resources by the individual. The approach of this study was therefore based on the capacity concept, referred to in the operational definition. At some point (or *red zone*) that perceived level of workload would be considered unacceptable and, therefore, a maximum number of inputs/activities could be identified. This would allow a specification as to the number of trains or amount of track or other parameters individual signallers could cope with. In doing this, there is a move towards a set of quantitative criteria and an analytical model, so that system designs that provide appropriate workloads could be defined.

This need to quantify workload and to be able to predict the expected workload at a specific train control centre was the desired outcome of the Spoornet study.

A number of problems were identified with the Pickup et al. (2005) approach, of which the most important were:

- Difficulties in distinguishing between inputs and outputs (e.g., is setting a route an influence on workload or an indicator of workload for a signaller?).
- Different qualitative impacts of different inputs on workload, e.g., setting routes or handling emergency phone calls.
- The need to account for individual differences in work strategy and workload perceptions.
- A large number of discrete and combined tasks that are completed as part of the signalling job.
- Change in performance shown, either through performance decrement or strategy change, before maximum workload capacity is reached.
- As systems move from being mechanical to computerised, there is less active, observable intervention; workload moves from the physical to the mental and is internalised to maintain situation awareness; and it becomes increasingly difficult to use observation measures.

This work (Pickup et al., 2005) provides a detailed account of MWL in signalling and follows with a presentation of a conceptual model. Primary elements in the model – such as loading factors, effort, demand and effects – are examined in detail. Finally, existing and proposed new workload measurement tools are introduced and positioned within the framework.

Mental workload in rail signalling was found to relate to the following elements:

- the number, complexity and interaction of tasks that are performed – over a period of time or at one point in time,
- the load as subjectively experienced – over a period time or at one point in time,
- the number of functions that have to be performed in different situations and scenarios, and
- the compatibility of working arrangements with the functions that need to be completed.

The main distinction is between workload as imposed by the system, somehow measured independently of individuals and their ratings of the ‘work’ that is ‘loaded’ on them (Hart and Staveland, 1988) and whether this reflects perceived task difficulty or level of effort exerted.

These multiple meanings, particularly as MWL is seen as an index that measures what people have to do and how they feel about it, have been the cause of great confusion in the rail industry. This is compounded by the fact that the concept of workload has a highly intuitive meaning for most people. Thus, the context in which mental workload has to be understood usually directs the interpretation of its meaning and the choice of assessment methods.

The multiple views of the term mental workload in the rail industry have directed the focus of this study to consider workload in a functional context. An understanding of signaller function was gained by reviewing goal and task analyses carried out as part of previous projects and by carrying out new interviews and observations as described earlier. From the task analyses, interviews and observations, it was suggested that signallers at the highest level have three main functions (See Figure 2.13).

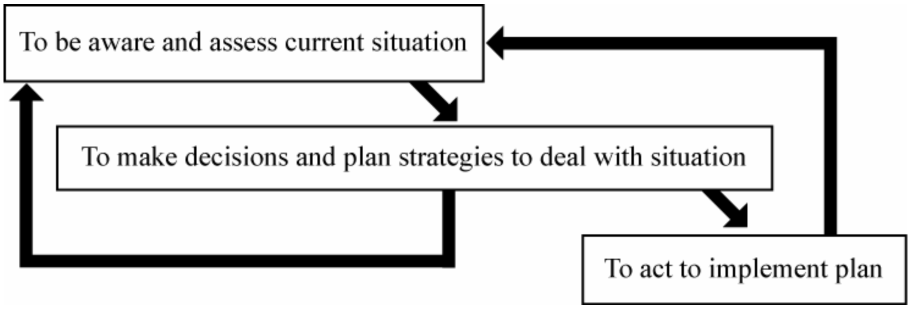


Figure 2.13: High level signaller functions (Pickup et al., 2005)

Workload tools previously used with rail signallers have included NASA TLX, timeline analysis and the Air Force Flight Test Centre (AFFTC) workload assessment scale. The interviews with rail human factors specialists and the personal experiences of those involved in the UK study indicate that these tools are not able to capture all aspects that can be considered relevant to workload of signallers. Often they are not appropriate for use in the field with civilian populations. Many tools only assess the workload of a single task (e.g., the NASA TLX, the Bedford scale and SWAT) whereas the studies of signallers and their environment suggest that multiple tasks are regularly completed concurrently. Many traditional workload assessment tools like SWAT only aim to assess dimensions that are related to high workloads (Reid and Nygren, 1988). The range of tools to be adopted by the rail industry should be capable of identifying situations of both underload (or at least low load) and overload, as both can impact upon health and safety and productivity. According to Pickup et al. (2005) the NASA-TLX is probably the most widely used workload measurement scale and extremely valuable in many situations. However, it was found to be impossible to use in real time in the field. Added to this, the wording appears to be a little obtuse for civilian and European audiences. More critically, it is not fully representative of the influencing variables most relevant to railway signallers.

These findings confirm the conclusion that was reached regarding existing tools and their appropriateness for use in the Spoornet context. This gave rise to the need to develop a customised tool for Spoornet.

Pickup et al. (2005) considered the development of a conceptual framework comprising the relevant dimensions of workload for the rail industry (and especially signalling) as essential support for a practical approach to workload measurement that may be usable in real situations. This conceptual framework has been developed gradually, expanded and refined in light of the relevant literature and also – critically – the authors' own studies of and with signallers. As a consequence, the evolving framework has been determined through a synthesis of empirical findings and theoretical interpretations and this synthesis is reflected in how the framework is explained.

The first simple conceptual framework of mental workload was based on the model proposed by Jahns (1973), which included input load, effort and performance. The starting framework reflected the effects of mental workload on both system performance and the wellbeing of signallers (Figure 2.14).

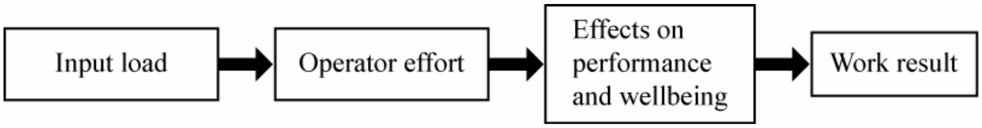


Figure 2.14: Adapted from Jahns' (1973) model (In Pickup et al., 2005)

The revised conceptual framework is shown in Figure 2.15 (below). Pickup et al. (2005) emphasise that this framework is not necessarily an operational model. It is proposed in order to develop and position a toolkit of methods with which to understand and assess mental workload. Thus, it explains the routes to measurement rather than the mechanisms that contribute to MWL. The framework also reflects the very different ways in which workload is conceived of and, therefore, measured by different investigators.

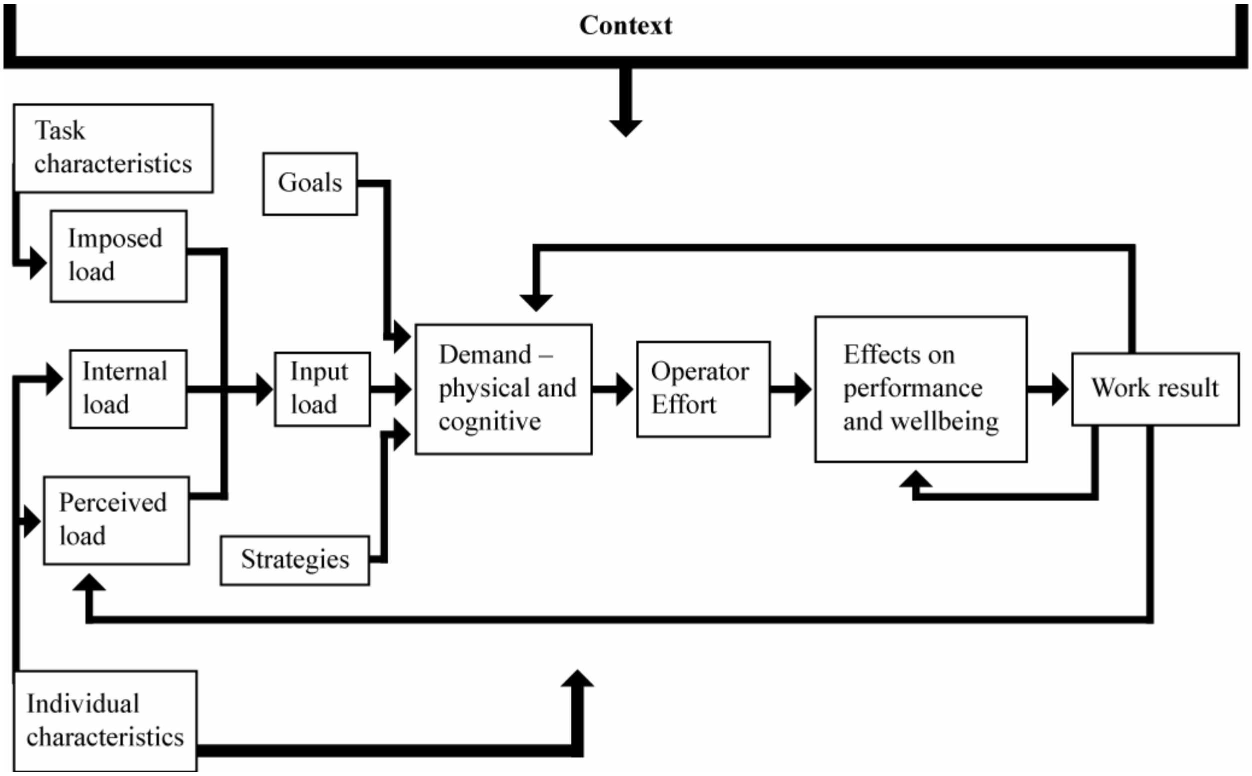


Figure 2.15: Developed conceptual framework of mental workload (Pickup et al., 2005)

The central elements of the framework are load, demand, effort and effects. Load is defined in terms of it being imposed and perceived. The interview analyses suggest different types of loading factors, that are categorised under several headings, for instance, organisation, environment (internal and external), situation (normal, abnormal or emergency working), team and job design, equipment, social situation, management and feedback on performance. There was considerable discussion during this study as to whether and how to represent context and individual differences within the framework. There is agreement that both context and individual differences may moderate the operation of the framework. However, questions such as whether imposed load and goals should be considered to be a part of the context, and whether experience, age, physical and mental wellbeing are part of individual differences or of context are still being debated. Goals and strategies were added as important intervening factors.

The analyses indicated that signallers viewed effort as a consequence of the demand created by loading factors. The perceived level of effort necessary to accommodate demands was related to:

- complexity of decision making required and options available,
- time sharing of activities,
- predictability of the situation,
- perception of the level of control within a situation,
- interruptions and distractions from the main activity,
- perception of likelihood of achieving intended goals (in relation to time, experience, organisational or environmental constraints),
- consequences of actions on the safety or performance of the system, and
- likelihood of being blamed for actions and existence of a blame culture.

The broader needs of workload measurement in the UK rail industry, especially those of Network Rail that were required for signalling, were referred to earlier. Obviously, no one tool would meet all those needs. For instance, assessing the number of functions that should be performed in a particular context requires a different form of measurement to that which would be required for assessing the load experienced at points in time by the person performing those functions. These different needs and the different understandings of the notion of workload have led to the conceptual framework embracing facets of load, goals and strategies, demands, effort and effects. The care taken to distinguish these notions means that care should also be taken to define measurement approaches and methods for each. In addition, various criteria against which to judge the value of workload tools have been defined by rail industry clients to include: allow assessment of acute load and also load over time; be diagnostic; be predictive; allow tracking of peaks and troughs of load; support direct assessment and be analytical; and reveal qualities of the performance. These are all in addition to the normal requirements for validity, reliability, acceptability, sensitivity, etc. No one measure will meet more than a few of these criteria. Understandably, there is no need to invent new methods and tools where existing ones will do. One use of the conceptual framework has been to clearly distinguish the purpose and focus of existing workload tools, to clarify what is available, even if it has been used and sometimes validated in other contexts. The conceptual framework also gave structure to the assessment of the potential value of the tools - for rail and for field, as well as for simulator use - undertaken with a sample by rail human factors specialists. The outcome of this exercise suggested that many tools currently available did not meet the needs of rail signalling, some had been used but with alterations in wording or method of administration, and one or two had potential but would need to go through trials and possible adaptation.

In terms of contextualising the Spoornet operational safety challenge, and in terms of appreciating the problem this research attempts to confront and resolve, the recent work undertaken in the railway environment and discussed here certainly provides a better understanding of the specific requirements of this industry. It also provides insight into the reasons why other well-known ‘industry-standard’ MWL tools are not useful in the railway context.

CHAPTER 3

METHODOLOGY

This chapter provides a detailed account of the development of the Mental Workload Index, which is the essence of this study. The chosen development approach and process is motivated and discussed after which a description of the specific steps that were followed to identify the factors to be included in the index and the allocation of weights to those factors, are detailed.

3.1 DEVELOPMENT APPROACH

The guiding principle in the approach was to develop a methodology that would provide an objective measurement of the 'pure', task-related mental workload factors that may be imposed on TCOs at specific train control centres. The methodology deliberately steers clear of assessing the performance of operators or the workload as subjectively experienced by specific train control operators.

However, because tasks are executed by human operators, it is difficult to isolate the individual from the task. Every attempt was made to adhere to the premise of identifying and measuring the task-related factors in the research design and in the development of the mental workload index.

A combination of quantitative and qualitative techniques was used in order to achieve the objectives of the study. The techniques include psychological and operational research techniques and physiological measurements.

For the proposed methodology to be successfully implemented and accepted by all relevant role players at Spoornet it was essential that they all formed part of the development process.

3.2 RATIONALE FOR THE METHODOLOGY

The model of Meshkati (1988) provides no guidelines as to which specific measurement techniques should be used for validation of mental workload in a context similar to that of this study. This lack of clarity as well as the lack of information on tasks and simulated environments in existing research, as confirmed by Cilliers (1992), as well as the conflicting research results referred to earlier, has lead to a problematic situation in the rail operational environment. The uncertainty and therefore the risk of using measurement techniques as described in the literature and then applying them to the context of rail safety, where operational decisions would be based on these measurements, is too high. It has therefore become necessary to mitigate the existing risk by exploring new ways of measuring mental workload.

It was decided that in the absence of a precedent or clear guidelines for a valid and reliable measurement technique, a new technique should be developed which will meet the criteria as stated in the literature. Considering the operational environment in which the results of this study may be applied, it was decided to pursue a participative and transparent development process that would allow for buy-in and high acceptance by the users of the system. These conditions were important as it will be operational managers, not mental workload specialists, who will use the tool to assist them to make operational decisions that relate to human factors and human performance.

The following were the criteria to be met by the Mental Work Load Index (MWLI) measurement method:

Construct validity was especially important because the process had to actually measure the factor *mental workload* and not another factor such as fatigue, which could be incorrectly perceived to be mental workload.

The predictive validity of the measurement technique for mental workload was also important as the mental workload measurement technique had to assess and/or predict the stress that the mental workload of a particular section could create for TCOs controlling that section.

Other criteria which it was considered important for the methodology to comply with were:

- Non-intrusiveness: TCOs perform safety-critical tasks and the method used may not distract them or affect their performance while executing their tasks.
- Operator and user acceptance.
- Sensitivity: the method should detect changes or differences in the mental workload imposed.
- Implementation requirements and affordability: the method should be easy to learn and administer, be portable and once developed, should be inexpensive to implement and maintain.

Although the primary objective of this project was the objective measurement of mental workload, there are other important aspects that had to be considered and that were related to the methodology for developing such a measure. An important envisaged benefit of the methodology is that it will be a scientifically developed process that could be defended should a legal, industrial relations or operational safety dispute arise with regard to the mental workload of TCOs in train control centres.

It must be borne in mind that the envisaged methodology is primarily aimed at determining the content of tasks and the mental workload demands these may place on operators. The mental workload should not be confused with on-the-job performance of the operators, which is a function of mental and physical workload combined with factors such as motivation, alertness, physical health and mental wellbeing.

In summary, the aim of the envisaged methodology is to create a management tool that can be applied to classify different railway lines and sections in terms of the mental workload they will impose on operators. Such classification will in turn facilitate planning and the allocation of resources to ensure safe train control practices and operations.

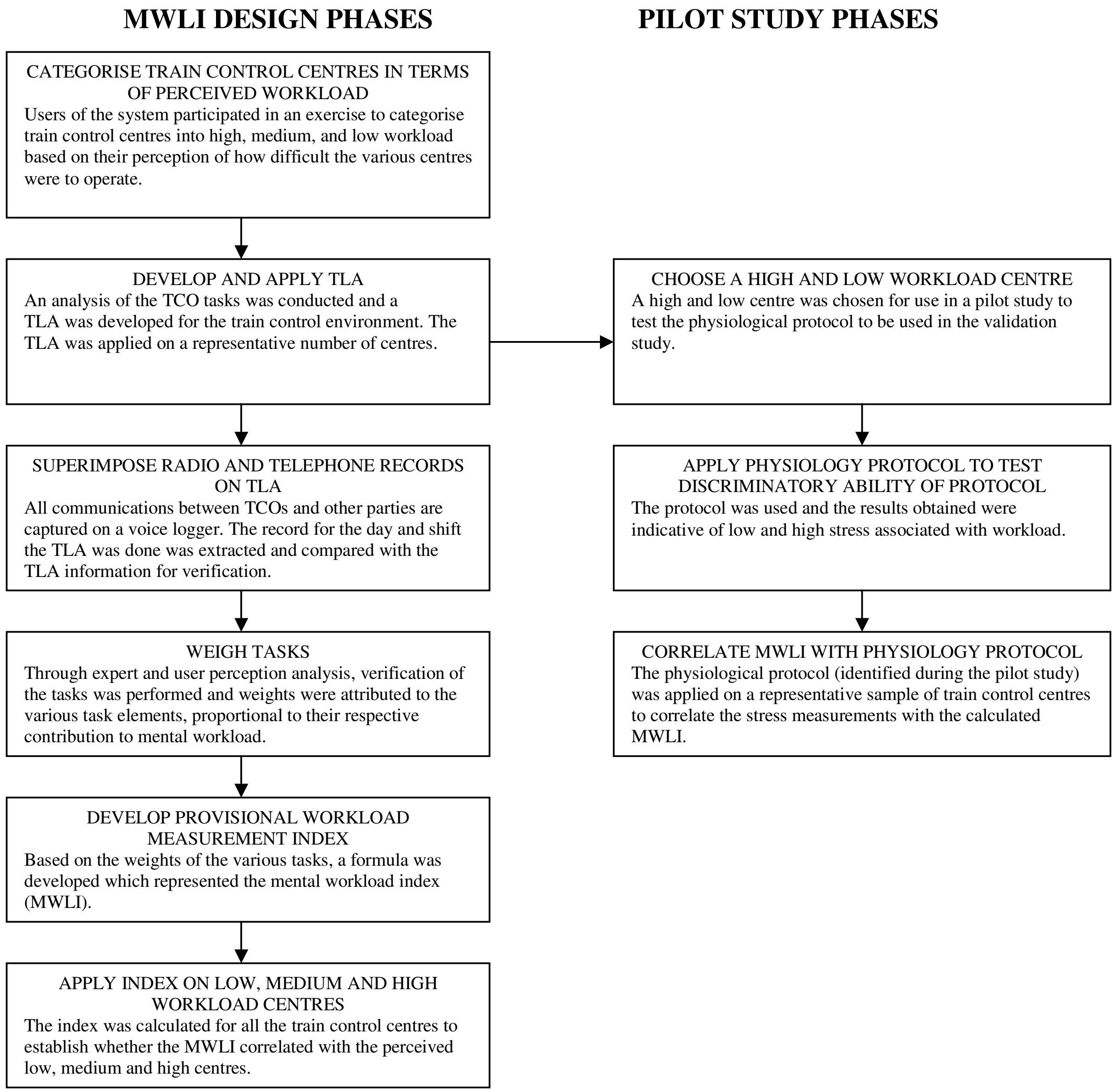
3.3 SUMMARY OF THE DEVELOPMENT PROCESS

The study was implemented in two phases, namely a project phase and a pilot study phase.

- The project phase commenced with a project proposal, which outlined the elements that would be addressed in the study and was based on a task analysis of TCO activities.
- The pilot study phase, which ran concurrently with the project phase, served as a feasibility study to determine whether some of the goals outlined in the proposal were attainable.

The following flow diagram outlines the development process of the mental workload index:

Table 3.1: Flow diagram of development process of the mental workload index



MWLI DESIGN PHASES

CATEGORISE TRAIN CONTROL CENTRES IN TERMS OF PERCEIVED WORKLOAD
Users of the system participated in an exercise to categorise train control centres into high, medium, and low workload based on their perception of how difficult the various centres were to operate.

DEVELOP AND APPLY TLA
An analysis of the TCO tasks was conducted and a TLA was developed for the train control environment. The TLA was applied on a representative number of centres.

SUPERIMPOSE RADIO AND TELEPHONE RECORDS ON TLA
All communications between TCOs and other parties are captured on a voice logger. The record for the day and shift the TLA was done was extracted and compared with the TLA information for verification.

WEIGH TASKS
Through expert and user perception analysis, verification of the tasks was performed and weights were attributed to the various task elements, proportional to their respective contribution to mental workload.

DEVELOP PROVISIONAL WORKLOAD MEASUREMENT INDEX
Based on the weights of the various tasks, a formula was developed which represented the mental workload index (MWLI).

APPLY INDEX ON LOW, MEDIUM AND HIGH WORKLOAD CENTRES
The index was calculated for all the train control centres to establish whether the MWLI correlated with the perceived low, medium and high centres.

PILOT STUDY PHASES

CHOOSE A HIGH AND LOW WORKLOAD CENTRE
A high and low centre was chosen for use in a pilot study to test the physiological protocol to be used in the validation study.

APPLY PHYSIOLOGY PROTOCOL TO TEST DISCRIMINATORY ABILITY OF PROTOCOL
The protocol was used and the results obtained were indicative of low and high stress associated with workload.

CORRELATE MWLI WITH PHYSIOLOGY PROTOCOL
The physiological protocol (identified during the pilot study) was applied on a representative sample of train control centres to correlate the stress measurements with the calculated MWLI.

3.4 STEPS IN THE DEVELOPMENT PROCESS AND ASSOCIATED INSTRUMENTS

The above development process comprised the following steps and related instruments:

A. A list of all RTO centres was compiled. Experienced representatives from the different groups of role players (expert users of the system) were requested to categorise the train control centres (and the sections that they control) into low, medium, and high in terms of traffic density, i.e., number of trains, maintenance activities and communication. Their instruction was to categorise the list of RTO centres into high, medium and low centres in terms of their perception of how difficult these different centres are to operate.

B. An analysis of the TCOs' tasks was conducted to determine the elements that should be included in the TLA.

C. A TLA checklist was developed specifically for the train control environment (See Addendum B).

D. Concurrent with step C, a pilot study was conducted to establish the physiological test protocol to be used in the validation of the mental workload index.

E. The TLA was applied per centre, using representative shifts, in order to determine what constituted the tasks, how much time was spent on the different activities, and to compare the possible differences between the different categories of centres.

F. Radio and telephone records were obtained and superimposed on the TLA. All communications between TCOs and other parties were logged on a voice logger. The record for the specific day and shift when the TLA was done was extracted from the voice logger and compared with the data captured on the TLA. The number and duration of radio and telephone communications were of specific importance because of their impact on the total workload experienced during a shift. It was, therefore, important that this information was captured correctly, and the voice logger records were used to verify the TLA information.

G. The various factors were weighted through a process of multicriteria decision modelling and expert analysis. This complex process comprised several steps, which in summary entailed the following:

- A group of expert users of RTO confirmed the elements of the TLA.

- The tasks/factors that contributed to workload, as identified through the application of the TLA, were weighted by a panel of user experts through a process of user perception analysis. (The Visual Interactive Sensitivity Analysis [VISA] was used to facilitate this process). This process was facilitated by an experienced operational researcher, who acted as an independent facilitator.
- Scales (Likert and quantitative) were developed to measure these factors. These scales are multidimensional in nature and they had to reproduce the proportionality of the specific factor, since the different factors do not contribute to the total workload on an equal basis.

H. The provisional workload measurement index was developed.

I. The provisional method was applied on the categories of low, medium and high workload centres (as categorised by expert users) to test the discriminatory ability of the method.

J. A study was undertaken to correlate physiological measurements of stress with the MWLI. This consisted of the following steps:

- Calculate the MWLI for a particular shift at selected train control centres.
- Apply the physiological protocol (as identified during the pilot study) on TCOs at the selected train control centres to determine what workload stress the subjects experienced and to correlate this with psychophysiological data for that same period.
- Correlate stress measurements and the calculated MWLI.

3.5 DETAILED DESCRIPTION OF THE DEVELOPMENT PROCESS

3.5.1 Listing and Categorisation of Radio Train Order/Track Warrant Train Control Centres

At the inception of the project, all train control centres that used radio train orders/track, warrant had to be identified in order to determine the extent of the study.

These centres were identified by the technical expert on the project team, who also categorised the centres into groups of high, medium and low intensity, based on the traffic volumes at the time. Expert users of the radio train order system, TCOs, and their supervisors were used to verify this categorisation. The rationale behind the categorisation by users of the system was to obtain an uncontaminated response based on the experience of the TCOs. At a later stage, this rating was correlated with the actual calculated MWL index.

3.5.2 Task Analysis

A task analysis was obtained to ascertain what the different elements are that constitute TCOs' tasks, and especially what contributes to their workload.

A task analysis gives an indication of what the task elements (activities) are that TCOs perform during a shift, but does not give an indication of how much time is devoted to these different activities. It must be kept in mind that the availability, or non-availability, of time, to perform work is the major component of mental workload and stress in the work context.

A task analysis, by means of task observation of the train control centre at the ZASM Building, Pretoria, was undertaken. These TCOs control the Rustenburg/Thabazimbi line. The following tasks were identified upon analysis of the information obtained:

- planning shift activities
- establishing which trains are scheduled for the shift
- plotting train movements on train schedule sheets
- controlling train movements by:
    - communicating by radio with train drivers
    - monitoring conversations of maintenance teams and other users of the track on the section on radio(s)
    - receiving requests for authorisations from train drivers by radio
    - planning the train movement on the train diagram
    - comparing the information on kilometre points on the section (on a different

  database) with planned movement to verify safety of the movement
  - completing the train order by referring to the train diagram in order to determine other train movements, establishing whether it is safe for the train to proceed to the next point, and obtaining an authorisation number from the records
  - writing the authorisation number, train number and train-driver details on a separate sheet
  - communicating the authorisation number to the train driver and waiting for confirmation from the train driver of the number
  - plotting train movement that have been authorised on the train diagram
- answering telephone calls from locomotive depot staff, maintenance staff or commercial personnel on train movements
- writing up reports regarding activities en route, such as locomotive failures, maintenance activities, shunting activities
- capturing incidents (locomotive failures or infrastructure problems, such as broken rails and criminal activity) and information on estimated times of arrival (ETAs) and estimated times of departure (ETDs) by accessing the data system on computers designed for capturing this data
- performing personal activities during discretionary time (making coffee, going to bathroom, reading)

In order of importance, the following are the priority tasks that are performed by TCOs:

- planning and authorisation of a train movement
- communicating the authorisation to the train driver
- ensuring that the authorisation was received correctly by the train driver
- keeping track of other activities, responding to unplanned events on the section, and communicating with the relevant parties

Devoe (1974) identified six functions that appeared to cover all the tasks of a dispatcher:

- prepare documentation
- conduct preliminary planning
- monitor/coordinate train movements

- initiate/stop train movements
- respond to unplanned events
- respond to emergencies

Of these, the functions of planning and decision making were considered the most critical.

Popkin et al. (2001) identified the following dispatcher tasks:

- planning
- controlling track use
- managing unplanned and emergency events
- record keeping and report writing

A comparison of the three sets of activities (Devoe, 1974, Popkin et al., 2001, and the Spoornet TCO tasks) shows significant congruency in terms of the planning and controlling/monitoring of train movements and responding to unplanned events, which, together with communication, appear to be the most critical tasks.

It is interesting that Popkin et al. (2001) note that as much as 75% of the dispatchers' shift may be spent communicating on the radio or telephone. This leaves little time for accomplishing other duties. Conversations with TCOs during development of the MWLI and during data collection indicated that sources of stress were related to telephone calls that caused them to divert their attention from train movements to other matters such as permanent way maintenance, requests related to the ETA/ETD of trains and general enquiries, as well as personal calls.

While performing the task observation it became evident that these tasks are often executed concurrently, such as receiving a request to authorise a movement while planning another movement or receiving a phone call while communicating an authorisation over the radio.

Once it was established which tasks occupied TCOs, it was necessary to determine how much time was spent on each activity. The Timeline Analysis was well suited for this purpose.

3.5.3 Timeline Analysis

A common approach to absolute workload and performance prediction is Timeline Analysis (TLA), which enables the system designer to 'profile' the workload that operators encounter during a typical mission, such as landing an aircraft or starting up a power-generating plant.

TLA is a process of systematically analysing how much time is devoted to which activities over a given period of time. It examines the temporal relationships among tasks and the duration of the tasks (Meister, 1985). As stated before, the time pressure factor is an important aspect in the total process of mental workload assessment. Timeline analyses are not only useful in determining workload but are also important functional breakdown techniques that can determine what the different elements of the task are and the time distribution over these elements. TLA allows observers to determine how much time is devoted to particular activities over a period of time, e.g., a shift.

The TLA technique provides for an objective mechanistic breakdown of the demands made upon operators in terms of time. It cannot measure the actual load placed on operators and therefore could not be used as a stand-alone method in the assessment of mental workload. The TLA is important in mental workload analysis for a number of reasons, e.g.:

- It forces analysts to break a job down into all its different tasks and activities.
- It gives an indication of how much time is devoted to a particular activity/task.
- It gives an indication of how much time in total has been taken up during a certain period of work.
- The most important value of a TLA is that it gives an indication of tasks that simultaneously lay claim to the attention of the operators.

It is these simultaneous demands on the attention of the TCOs that contribute substantially to the TCOs' workload and that could result in stress and contribute to judgement errors. TLA is critical in workload analysis, since the time factor is an important aspect in the process of mental workload assessment.

A TLA technique that was used effectively in the dispatching environment is a modified version of the Task Analysis Workload Measure (TAWL) by Popkin et al., (2001). The TAWL, developed for the assessment of military helicopter flight crews, was modified to suit the dispatching operational environment and is referred to as the mTAWL.

The mTAWL is a task-oriented approach that assumes that dispatchers perform multiple, simultaneous tasks through time, and that these individual tasks may vary in their demand on the dispatchers' performance resources. The mTAWL treats workload as the sum of the difficulty of all concurrent tasks for each minute of an observation. Two dispatchers may handle an equal number of trains, yet significant differences may exist in workload across a shift if one of the dispatcher's activities takes place within a short period of time while the other dispatcher's load is spread evenly over time. The mTAWL is sensitive to this difference. The mTAWL also takes the difficulty of a task into account. Two tasks may be of equal time duration, yet one task may call on more resources of listening, watching, thinking or overtly acting than the other. The mTAWL refers to these resources as auditory, visual, cognitive or psychomotor channels. The mTAWL method calculates workload by adding the loads for each of the individual channels across all tasks for each minute of a dispatcher's shift.

A TLA format was used in this study during the first phase of the MWLI development (see Appendix B). The purpose was to record the activities of TCOs at particular train control centres. An eight-hour shift was divided into 15-minute sections and the activities performed (often concurrently) were captured per 15-minute periods. The tasks identified using the task analysis and TLA were further analysed and were eventually included in the Mental Workload Index.

3.5.4 Pilot Study

The purpose of the pilot study was to develop and to evaluate the proposed physiological protocols and data collection procedures, and to determine whether the selected protocols would discriminate between high and low mental workload. The Department of Medical Physiology, University of Pretoria, assisted in determining which measures would be practicable in the specific circumstances for the selected protocol. The research was conducted *in situ*, in the TCOs' workplace. The assurance that the data collection would not interfere with the TCOs' ability to perform their job was a major consideration in the design of the study. In order to satisfy the requirements of the Ethics Committee of the University of Pretoria and to ensure objectivity and consistency in the data collection process, staff members of the Medical Physiology department were responsible for performing the physiological measurements.

With the assistance of the team of physiologists a protocol to measure allostatic load associated with mental workload was developed. The protocol for allostatic load consists of 13 measurements, most of which are measured in blood (such as total cholesterol, adrenalin and noradrenalin). The collection of blood samples was not considered feasible for this study due to its intrusive nature. The parameters included in this study were:

- Blood pressure – diastolic (DBP) and systolic (SBP) blood pressure were measured with a digital electronic blood pressure meter (ALP K2, model DS-125D, Japan).
- Heart rate – measured with Polar heart rate monitors (Polar Electro).
- Cortisol – free salivary cortisol was determined with a Salivary Cortisol ELISA kit (SLV-2930, DRG Instruments GmbH, Frauenbergstrasse, Marburg, Germany).
- The Body Mass Index (BMI) was also calculated as it provides an indication of the metabolic effects of stress.
- TLA.
- Cohen's Perceived Stress Questionnaire.

Permission was obtained from the Ethics Committee of the University of Pretoria for the physiologists at the Department of Medical Physiology to use these protocols, and a pilot study was performed at the Welgedag train control centre during June 2001. The reasons for choosing the Welgedag train control centre were twofold:

- This specific train control centre controls more than one section; the one section was rated as high and the other as low in terms of traffic density and therefore perceived workload.
- Geographically the centre was within reasonable travelling distance from Pretoria should repeat visits become necessary.

The results were conclusive and indicated that the measurements were sensitive enough to distinguish between high and low workloads (see Appendix C for the Pilot study report).

The following problems were, however, encountered:

- Electromagnetic interference was experienced in the train control centre, which affected the reliability of the Polar heart rate monitor (a chest strap which transmits a radio signal to a wrist receiver). It was recommended that a different type of monitor be used that makes use of a direct cable connection rather than the radio transmission of a signal.
- Because it was subsequently established through a literature review that most of the research indicated that heart-rate variability is a more sensitive measure of the stress associated with mental workload, and since a new instrument, in any event, needed to be sourced to replace the Polar heart rate monitor, it was decided to measure heart-rate variability (HRV) rather than heart rate. Using heart-rate variability would put this study on an equal footing with international studies and would make comparative analysis easier.

3.5.5 Application of the Timeline Analysis (TLA)

Using representative shifts, the TLA checklist was applied to a number of train control centres. The aim was to determine whether the activities in all train control centres were the same, and to develop a better understanding of the time spent on each activity (which could give an indication of the complexity of the specific activity). This was necessary when developing the index and when assigning weights to different activities. The TLA was also used to compare high- and low-traffic-density centres with one another.

3.5.6 Radio and Telephone Records

Radio and telephone records for all the identified centres were obtained and superimposed on the TLA. In other words, this information was added to the TLA information to provide a clearer picture of the tasks and the time spent on them.

3.6 DEVELOPMENT OF THE MENTAL WORKLOAD INDEX

Considering the wealth of information on mental workload assessment that has accumulated over the last three decades, the question, “Why a new methodology?” can legitimately be posed.

The need was identified for a fit-for-purpose and easy-to-use means of assessing and predicting the potential workload that could potentially be experienced at train control centres. The MWLI does not claim to replace or to be superior to any of the existing assessment tools, especially as it does not measure the actual workload that can be experienced by TCOs at any specific point in time. It is based on task-load factors and on a number of mediating factors. It is aimed at providing rail operations managers with a means of identifying and highlighting potential problem train control centres, without their needing specialised assistance to do so. The MWLI could be calculated as a desk-top exercise because it uses objectively collected information that can be obtained from a central database. It does not necessitate an on-site review of the activities at a specific train control centre. The tool is also useful in developing different scenarios by changing the values of the factors and then predicting what the workload will be. The value of this is that workload estimates could be determined before changes to train control systems or personnel are made.

An important consideration in the development of a tool to be used in the train control environment was that of intrusiveness. As discussed in an earlier section, intrusiveness refers to the tendency of a measurement technique to cause unintended degradations in ongoing primary-task performance. This can pose potentially serious problems in the application of a workload measurement technique. Such problems are primarily related to the interpretation of results obtained with an assessment procedure, and with the application of techniques to operational environments. Levels of intrusiveness that may be acceptable in the laboratory might not be tolerable in operational environments where any compromise in system safety would be unacceptable.

Many of the measurement procedures that have been developed are very promising, but they are still largely restricted to research environments. It appears equally clear that because of the complexity of the workload construct, it is unlikely that any single measure will be completely adequate in providing the type of applied measurement mechanism that is desired, and at the same time be applicable to all kinds of applied work situations (Reid and Nygren, 1988).

The Mental Workload Index was developed over a number of iterations and after several group sessions. The group sessions comprised the same group of people every time, and were facilitated by an independent operational researcher. The group represented all potential stakeholders namely technical experts, users of the radio train order system (TCOs) and their supervisors, as well as trade union representatives. The result was an evolutionary developmental process of the mental workload index.

The outcome of the iterative process was consensus and acceptance of the following factors to be included in the MWLI (See summary in Table 3.2):

- Three task elements (additive), descriptive of the content of work – the sum of the elements makes up the *quantity* of work. The three selected task elements are the number of data transactions, the number of authorisations, and the number of communications via telephone and radio.
- Eleven moderating factors (multiplicative), descriptive of the *nature* of the work – the moderators influence the complexity of the work, irrespective of the quantity, e.g., the type of shift and the experience of the TCOs.

The proposed model was initially of the form:

$$I = \sum_{i=1}^{3} T_i * \prod_{j=1}^{11} M_j$$

MWLI = (sum of task elements) × (product of moderating factors) or
where I is the MWLI in arbitrary units, T_i, i=1,2,3 are the three task elements, and M_j, j=1, 2,…,11 are the maximum values of the eleven moderating factors.

Due to the difference in the relative importance of each task element and moderating factor to the overall mental workload, both the task elements and the moderators were weighted.

The weighting factors for each were also determined by the work group. The MWLI was refined as follows:

MWLI = (sum of weighted task elements) × (product of weighted moderating factors) or

$$I = \sum_{i=1}^{3} T_i * W_i * \prod_{j=1}^{11} M_j * V_j ,$$

where I is the MWLI, W_i, are the weights of the three task elements T_i, i.e., i=1,2,3, and V_j, are the weights of the eleven moderating factors M_j, with j=1, 2,…,11, and $\sum$, $\prod$ designate sum and product over the indicated terms.

Table 3.2: Elements of the MWLI

Task Elements		Moderators		
Item no.	**Description**	**Item no.**	**Description**	**Maximum Weight (%)**
1	Number of data transactions	1	Shift	12
		2	Experience	18
2	Number of authorisations	3	Interface complexity	5
		4	Running times	8
3	Number of communications via telephone/radio	5	Crossing places	6
		6	Platform location	3
		7	Trains vs. number of crossing places	11
		8	Type and mix of trains	9
		9	Locomotive depots	10
		10	Shunting activities	14
		11	Topography	4

The range of the MWLI is from 0 to NN, where the minimum corresponds to zero transactions, authorisations and communications, and the maximum NN corresponds to the peak number of transactions, authorisations and communications, with the worst case for each of the moderators. In practice the MWLI ranged from 89 to 5789 for the 36 train control centres that were part of the study.

3.7 IDENTIFICATION OF TASK AND MODERATING FACTORS

An initial task analysis of the activities performed by TCOs during a shift was performed in order to identify the tasks. During a TLA the activities and number of times they are performed, per operator, are captured every 15 minutes. Several activities directly related to the task were identified. The project team, together with expert users of the system, analysed the data obtained from the TLAs.

The following iterations were performed in order to reach final consensus and acceptance of the factors to be included in the MWL Index.

Table 3.3 Iterative process in the development of the mental workload index

	Task Elements	**Moderators**
First Iteration	• No. of authorisations (no. of trains × no. of radio authorisations per train) • Weighted no. of authorisations (no. of authorisations × authorisation weight – weight as yet undetermined – arbitrary figures used to illustrate concept) • Total no. of actions (weighted no. of authorisations + no. of telephone calls)	• Difficulty of the section • Shift type • Tiredness • Error impact • Experience (Moderators as yet unquantified – arbitrary figures used to illustrate concept)
	MWL Index = Sum of actions (weighted) × Product of moderators	
Second Iteration	• No. of data transactions • Weighted no. of data transactions (no. of data transactions × data transaction weight – weight as yet undetermined - arbitrary figures used to illustrate concept) • No. of radio authorisations • Weighted no. of authorisations (no. of radio authorisations × authorisation weight – weight as yet undetermined – arbitrary figures used to illustrate concept) • No. of telephone/radio communications • Weighted no. of actions (weighted no. of data transactions + weighted no. of authorisations + no. of telephone/radio communications)	• Inherent difficulty of the section • Scheduling complexity • Interfacing difficulty • Type/mix of trains • Type of shift (fatigue) • Duration of shift • Experience (Moderators as yet

	Task Elements	**Moderators**
		unquantified – arbitrary figures used to illustrate concept)
	MWL index = sum of actions (weighted) × product of moderators	
Third Iteration	• No. of data transactions • No. of authorisations • Weighted no. of authorisations (no. of authorisations × authorisation weight) – authorisation weight = 15 (see discussion below) • No. of other telephone/radio communications • Weighted no. of communications (no. of other telephone/radio communications × telephone/radio communications weight) – telephone/radio communications weight = 5 (see discussion in paragraph 4.4.1) • Weighted no. of actions (no. of data transactions + weighted no. of authorisations + weighted no. of communications)	• Inherent difficulty of the section • Planning complexity • Interface complexity • Type/mix of trains • Type of shift (fatigue) • Duration of shift • Experience (Moderators as yet unquantified - arbitrary figures used to illustrate concept)
	MWL index = sum of actions (weighted) × product of moderators	
Fourth Iteration	• No. of data transactions • No. of authorisations • Weighted no. of authorisations (no. of authorisations × authorisation weight) – authorisation weight = 15 (see discussion in paragraph 4.3.1) • No. of other telephone/radio communications • Weighted no. of communications (no. of other telephone/radio communications × telephone/radio communications weight) – telephone/radio communications weight = 5 (see discussion in paragraph 4.4.1) • Weighted no. of actions (no. of data transactions + weighted no. of authorisations + weighted no. of communications)	• Shift (weight: 12%) • Experience (weight: 18%) • Interface complexity (weight: 5%) • Running times (weight: 8%) • Crossing places (weight: 6%) • Platform location (weight: 3%) • Trains vs. crossings (weight: 11%) • Type & mix of

	Task Elements	Moderators
		trains (weight: 9%) • Loco depots (weight: 10%) • Shunting (weight: 14%) • Topography (weight: 4%) (Discussion of weight allocation in paragraph 4.4.2)
	Workload index = total no. of actions (weighted) × product of moderators (weighted)	

The process of identifying task elements and moderators and determining the weight of each spanned a period of eight months.

3.8 DEFINITION OF TASK AND MODERATING FACTORS

3.8.1 Definition of Final Task Elements

3.8.1.1 Number of data transactions

This factor includes data that TCOs have to capture over and above the data required for an authorisation. Data transactions that are typically executed are Spoornet Information Management System (SIMS), estimated time of arrival (ETA), and estimated time of departure (ETD) entries. Data included in the calculation of the Mental Workload Index represents the capturing of SIMS and ETA/ETD data. No other data is captured by TCOs. Some train control centres do not capture any data transactions. This is indicated by a 0 (zero) on the table. The total number of data transactions is calculated by adding the number of SIMS and ETA/ETD transactions executed.

The rationale for including this factor was that any activity that would increase the workload of TCOs had to be included in the index. This activity increases the time demands on TCOs and requires accuracy and therefore concentration.

The number of data transactions is captured on a central database for SIMS and the information was extracted directly from that database.

3.8.1.2 Number of Authorisations

This factor represents the essence of the output of TCOs. Adding the number of authorisations for the purposes of calculating the mental workload index is a simple exercise and belies the complexity of executing the task.

The issuing of an authorisation for a train to proceed to a point further in the section entails the following actions:

- The TCOs have to consult the train diagram. The train diagram contains the planned train movements for the shift as well as the actual train movements. Planning is done by the TCOs at the start of the shift and actual movements are updated as trains move through the section. This type of train control does not have a panel that shows the movement of trains in the section. The TCOs have only a hand-drawn train diagram (computer-based if TWS is used in conjunction with RTO) that shows train movements and where trains will cross.
- Trains crossing each other in the section require careful planning as this type of train control occurs on a single track, and crossings can only be done at stations or loops in the section.
- Before a train can be authorised to proceed to a predetermined point, the TCOs need to ensure that traffic from the opposite direction is on time as indicated by the train diagram. If not, alternative plans have to be made, such as using an alternative crossing place or keeping a train (in most instances the train that is behind schedule) back. The train diagram has to be adjusted to reflect this change.
- All authorisations occur via radio communication. The train driver contacts the TCO on the radio when the point up to where authorisation was received has been reached. The TCO then executes the above-mentioned activities before authorising the train driver to proceed. The TCO radios the train driver when the planning for the movement has been completed. The train driver performs identity verification by providing the train number as well as the kilometre point where the train is waiting. The TCO

confirms this information over the radio while writing it down on an authorisation sheet. An authorisation number is given to the train driver which is in turn confirmed by the train driver.

- The train then departs from the waiting point. The TCO updates all the information on the train diagram and proceeds to plan for this particular train movement, i.e., when the train will reach the next point according to train running times, which are predetermined. In this study it was found that anything between 1 and 97 authorisations could be issued during an eight-hour shift.

The complexity of this element is captured in the moderating factors as well as in the weight assigned to this task (see explanation below of how weights were assigned).

The number of authorisations was obtained from the manual records kept by the TCOs.

3.8.1.3 Number of Telephone/Radio Communications

Communication via radio or telephone is an integral part of the job function of TCOs. All authorisations are communicated via radio or telephone. These are not the only communications TCOs deal with, and a variety of other requests and queries are also channelled through them, such as the activities of maintenance teams, information on ETA and ETD, information on breakdowns or locomotive failures, and shunting activities.

The information for this element was obtained from the voice logger, which records the communications of the TCOs during their shift. The number of communications was calculated by adding the individual communications captured on the recording system.

3.8.2 Definition of Final Moderators

It would of course be meaningless to attempt quantification of the moderators if all members of the work group did not use the same definitions. The definitions for the earlier versions of the model have not been included in this document. The final set of moderators was defined as follows:

3.8.2.1 Shift

Time of day and length of shift.

All possible shifts that are normally worked were captured in the scale, but it might happen that, due to special circumstances, such as people being on sick leave, TCOs would work a longer shift that does not fully correspond with the options provided. The option on the scale with the greatest overlap with the actual shift time will then be selected.

3.8.2.2 Experience

Level of applicable experience expressed in number of years.

Applicable means experience in the specific train control system under evaluation. Relevance of experience is reduced by interrupted service.

3.8.2.3 Interface complexity

Interface with other train control systems such as colour light signalling.

Interfaces can be single or multi-interface, meaning that in one direction of the section there could be an interface with colour light signals and in the other direction there could be semaphore signalling. TCOs have to take cognisance of these interfaces when communicating with train drivers.

3.8.2.4 Running times

Running times of trains in a specific section within limited variations. Running times are pre-established and are contained in rule books.

3.8.2.5 Crossing places

Types of points sets, e.g., hand points, electrical points or self-normalising points.

Critical gradients on approach to a crossing place and level crossings within a train length, from the facing points, are also important factors to consider.

3.8.2.6 Platform location

The presence or absence of platforms in the section.

This complicates the planning tasks of the TCOs. For example, if a passenger train is brought into a station that is simultaneously used as a crossing place with another train, and passengers need to embark or disembark, the train must be taken into the station on the side where the platform is located.

3.8.2.7 Number of trains vs. number of crossing places

The impact of the relationship between the number of trains versus the number of crossing places can have considerable impact on planning complexity. The correlation between the two variables is not linear.

3.8.2.8 Type and mix of trains

The presence of various types of trains with varying priorities.

Different trains carry different priorities. The Blue Train (a luxury passenger train) has the highest priority and always gets preference if one train has to wait for another at a crossing place. Name trains, such as the Trans Karoo, Diamond Express or Trans Natal have second priority. Hazardous materials trains, such as those that transport explosives or petrochemical products have third priority. Block trains, which are assigned to a specific client such as Sappi, have fourth priority. Other passenger and goods trains follow further down on the priority list. TCOs need to know which types of train run on their sections in order to plan properly.

3.8.2.9 Locomotive depots

The presence or absence of locomotive depots en route.

The presence of locomotive depots results in more train movements due to locomotives being shunted to couple or uncouple loads.

3.8.2.10 Shunting

The presence of shunting activities in the section.

3.8.2.11 Topography

The nature of the terrain and landscape through which the section runs.

Scales to describe these factors were developed. These scales are multidimensional and they reproduce the proportionality of the specific factor, since the different factors do not contribute to the total workload on an equivalent basis (Table 3.4).

Table 3.4: Definitions of moderators

Moderator	**Definition**	**Scale**	**Range**
Shift	Time of day and length of shift	8-hour time periods over a 24-hour period	1-5
Experience	Level of applicable experience expressed in number of years (*Applicable* means experience in the specific train control system under evaluation. Relevance of experience is reduced by interrupted service.)	Expressed in number of years	1-4
Interface complexity	Interface with other train control systems	Same system or different system; single- or multi-interface	1-4
Running times	Running times of trains in a specific section within limited variations	Uniform or varying running times (dependent on type of trains)	1-4
Crossing	Types of points sets, e.g., hand points,	Uniform or	1-2

places	electrical points, self-normalising points, etc. Also critical gradients on approach to crossing places and level crossings within a train length from the facing points	varying	
Platform location	The presence or absence of platforms in the section. This complicates the planning tasks of TCOs	Presence and position of platforms	1-4
Number of authorisations vs. number of crossing places	The impact on planning complexity of the relationship between number of authorisations vs. number of places where trains can cross	Low to very high impact – expressed in matrix	1-4
Type and mix of trains	The presence of various types of trains with varying priorities	Types of trains and variation	1-4
Locomotive depots	The presence or absence of locomotive depots en route	Presence and number and depots	1-4
Presence of shunting yards/ activities	The presence of shunting activities in the section	Presence of shunting yards, private sidings	1-4
Topography	The nature of the terrain and landscape through which the section runs	Flat through hilly; urban through city	1-4

3.9 ALLOCATION OF WEIGHTS AND QUANTIFICATION OF MODERATORS

3.9.1 Authorisation and Communication Weights

During a facilitated work session using a modified nominal group technique with the work group, the participants came to the conclusion that capturing data transactions was the simplest of the tasks carried out by TCOs.

The modified nominal group technique process was facilitated by an independent operational researcher and entailed eliciting ideas from participants, in this case their ideas about the essential tasks of TCOs. The aim was to facilitate open participation, where everyone’s ideas had equal importance, irrespective of their position or status. Ideas were written on post-it™ paper notes (one idea per note) and stuck on a blackboard. With the assistance of the facilitator, ideas with the same meaning were grouped together and the idea was then classified in terms of importance. Using their perceptions of the work content of a data transaction as a basis, the work group reached consensus that an authorisation transaction entailed fifteen times as much work (in terms of content and complexity) as a data transaction. Similarly a telephone or radio communication on average entailed five times as much work as a data transaction. Hence the weights used in the model:

Data transaction	: 1
Authorisation	: 15
Telephone/radio communication	: 5

3.9.2 Moderator Weights

Once again, during a series of sessions and using the modified nominal group technique, the work group decided that *platform location* was the least taxing of the moderating factors.

Experience of the operator was considered to be the factor that contributed most to the mental workload, and group consensus was that experience should carry six times the weight of *platform location*. The other moderating factors were then placed in order of diminishing importance between *experience* and *platform location*, and a simple geometric representation was used to assist the group in reaching a consensus on the relative contribution of each moderator.

The resulting values were then normalised (scaled so that they added up to 100%), resulting in the weights appearing in the Mental Workload Index.

3.9.3 Scale Points of Moderators

A similar logic to that used to allocate weights to the moderators was employed to determine scale points for the moderators.

Consider, for example, the moderator *shift type*. The work group first defined scale points 1 to 5. They then reached consensus that scale point 1 (i.e., the 06:00 to 14:00 shift) was the least taxing, and a value of 1.0 was allocated to this point (multiplication by 1.0 leaves the index unchanged, i.e., there is no moderation). Next, consensus was reached that scale point 5 (i.e., the 18:00 to 06:00 shift) was the most taxing, and it was allocated an arbitrary rating greater than 1. The decision of the work group relating to the most and least taxing shifts was verified with human factors and stress experts. It was confirmed that in terms of physiological factors (circadian rhythm, drowsiness, and mental alertness) the 18:00 to 06:00 shift was the most difficult and the 06:00 to 14:00 shift was the easiest. The other scale points were then allocated ratings between 1.0 and this maximum value. The intermediate ratings were debated and adjusted until the group agreed that the differences in allocated ratings were in the proper relation to their assessments of the differences in mental workload.

Another example is the factor *experience as a TCO on the particular system.* The work group identified four phases in the career of TCOs that could moderate experienced workload:

Table 3.5: Example of scale point development

Scale	Factor	Rationale
Up to 2 years after being licensed	2	Too little experience to effectively cope with difficulties and demands. Highest moderating impact.
2 to 7 years after being licensed	1	Suitably experienced to cope with problems, find solutions and remain in control. Remains vigilant and alert. No moderating effect.
7 to 12 years after being licensed	0.8	Experienced, has developed an intuitive approach to the demands – 'knows' when problems are likely to occur. Makes load easier.
More than 12 years after being licensed	1.2	Becomes complacent – not sufficiently alert and vigilant. Moderates workload – will increase workload

The allocated factors were based on the experience the work group had had in their own careers and with their subordinates, whom they manage daily.

Other scale points of the moderators were similarly determined to be the most or least taxing. In all the other instances, the expertise and experience of the work group and a consensus decision determined the scale point values.

Finally, when the same procedure had been repeated for all the moderators, all the ratings were mathematically scaled in such a way that the highest moderating value in each case corresponded to the weight allocated to the moderator. For example, the maximum moderating factor for *shift worked* (weight 12%) is 1.12 (the maximum moderating effect was achieved by multiplying by 1.12 which means that the index value was moderated by 12%). The same logic was applied to all the moderators resulting in the descriptions reflected in the table below:

Table 3.6: Description and weights of moderating factors

1. SHIFT TYPE

1.1 Shift Worked (Weight: 12%)

- Time of day and length of shift
- Select the time period with the greatest overlap with the actual shift time

Factor	Scale Pt	Description
1.00	1	06:00 – 14:00
1.04	2	14:00 – 22:00
1.12	3	22:00 – 06:00
1.10	4	06:00 – 18:00
1.15	5	18:00 – 06:00

2. EXPERIENCE AS A TCO ON THE PARTICULAR SYSTEM

2.1 Experience in Years on RTO/TWS (Weight: 18%)

- Level of applicable experience expressed in number of years
- *Applicable* means experience in the specific train control system under evaluation
- Relevance of experience is reduced by interrupted service

Factor	Scale Pt.	Description
2	1	Up to 2 yrs after being licensed
1	2	2 to 7 yrs after being licensed
0,8	3	7 to 12 yrs after being licensed
1,2	4	More than 12 yrs after being licensed

3. PLANNING COMPLEXITY (Total Weight: 30%)

3.1 Interface Complexity (Weight: 5%)

- Interface with other train control systems (single or multi-interface)

Factor	Scale Pt.	Description
1	1	Single interface, system to system
1,2	2	Multi-interface, system to system
1,4	3	Single interface, with another system
1,6	4	Multi-interface, with another system

3.2 Running Times between Crossing Places (Weight: 8%)

- Running times of trains in a specific section within limited variations

Factor	Scale Pt	Description
1	1	Uniform running times, long (>30 minutes) (Variations <30%)
1,8	2	Uniform running times, short (<30 minutes) (Variations <30%)
1,3	3	Varying running times, long (>30 minutes) (Variations >30%)
1,6	4	Varying running times, short (>30 minutes) (Variations >30%)

3.3 Types of Crossing Places (Weight: 6%)

- Types of points sets, e.g., hand points, electrical points, and self-normalising points. Also critical gradients on approaches to crossing places and level crossings within a train length from the facing points

Factor	Scale Pt	Description
1	1	Uniform crossing places (same types of points, etc.) (Variations <30%)
1,5	2	Varying crossing places (different types of points, etc.) (Variations >30%)

3.4 Location of Platforms (Weight: 3%)

Factor	Scale Pt	Description
1	1	No Platforms
1	2	All platforms on both lines
1,2	3	Less than 50% of crossing places have platforms on one line only
1,5	4	More than 50% of crossing places have platforms on one line only

3.5 Number of Authorisations per Shift vs. Number of Crossing Places (Weight: 11%)

- Impact of relationship of authorisations vs. crossings places on planning complexity (See Matrix – Figure 6.1)

Factor	Scale Pt	Description
1	1	Low impact according to matrix
1,2	2	Medium impact according to matrix
1,4	3	High impact according to matrix
1,5	4	Very high impact according to matrix

4. INHERENT DIFFICULTY OF THE SECTION (Total Weight: 40%)

4.1 Type/Mix of Trains (Weight: 9%)

- The presence of various types of trains with varying priorities

Factor	Scale Pt.	Description
1	1	One type of train (e.g. goods train) with similar priority
1,05	2	One type of train with varying priority (e.g. goods trains – hazmat/other materials)
1,15	3	More than one type of train with varying priority (e.g. goods trains, passenger trains – excluding Metro trains, hazmat train)
1,30	4	More than one type of train with varying priority (e.g. goods trains, passenger trains – including Metro trains, hazmat train)

4.2 Presence of Locomotive Depots (Weight: 10%)

- The presence or absence of locomotive depots en route

Factor	Scale Pt	Description
1	1	No locomotive depot
1,1	2	One locomotive depot (at start or end)
1,15	3	Locomotive depot at start and end
1,2	4	More than two locomotive depots (including depot(s) en route)

4.3 Presence of Shunting Yards/Activities (Weight: 14%)

Factor	Scale Pt	Description
1	1	No shunting en route or if shunting yard or private siding is at end or start of section
1,05	2	Shunting yards en route
1,25	3	Private sidings and/or shunting activities en route
1,30	4	Private sidings and/or shunting activities plus shunting yards en route

4.4 Topography (Weight: 4%)

- The nature of the terrain and landscape through which the section runs
  (The impact of the presence of tunnels was considered but the applicable operating principles have no significant impact on TCOs and it was therefore not considered relevant for the index)

Factor	Scale Pt.	Description
1	1	Karoo landscape; small towns few and far between
1,05	2	Flat terrain; high-density built-up environment
1,1	3	Hilly terrain; small towns few and far between
1,2	4	Hilly terrain; high-density built-up environment

Table 3.7, below, summarises the moderating factors in terms of their scale points, moderating weight and the moderating factor used in the calculation of the MWLI.

TABLE 3.7: Moderators – description of scale points and associated moderating factors

Moderator	Moderating weight	Scale Point	Description	Factor	Moderating factor
Shift worked	0.12	1	0600 – 1400	1	1.000
		2	1400 - 2200	1.04	1.032
		3	2200 - 0600	1.12	1.096
		4	0600 - 1800	1.1	1.080
		5	1800 - 0600	1.15	1.120
Experience in years on RTO/TWS	0.18	1	Up to 2 yrs after being licensed	2.5	1.180
		2	2 to 7 yrs after being licensed	1.25	1.030
		3	7 to 12 yrs after being licensed	1	1.000
		4	More than 12 yrs after being licensed	1.5	1.060
Interface complexity	0.05	1	Single interface, system to system	1	1.000
		2	Multi-interface, system to system	1.2	1.017
		3	Single interface, with another system	1.4	1.033
		4	Multi-interface, with another system	1.6	1.050
Running times between crossing places	0.08	1	Uniform running times, long (>30 minutes) (Variations <30%)	1	1.000
		2	Uniform running times, short (<30 minutes) (Variations <30%)	1.8	1.080
		3	Varying running times, long (>30 minutes) (Variations >30%)	1.3	1.030
		4	Varying running times, short (>30 minutes) (Variations >30%)	1.6	1.060
Types of crossing places	0.06	1	Uniform crossing places (same types of points, etc.) (Variations <30%)	1	1.000
		2	Varying crossing places (different types of points, etc.) (Variations >30%)	1.5	1.060
Location of platforms	0.03	1	No Platforms	1	1.000
		2	All platforms on both lines	1	1.000
		3	Less than 50% of crossing places have platforms on one line only	1.2	1.012
		4	More than 50% of crossing places have platforms on one line only	1.5	1.030

Moderator	Moderating weight	Scale Point	Description	Factor	Moderating factor
Number of authorisations per shift vs. number of crossing places	0.11	1	Low impact according to matrix	1	1.000
		2	Medium impact according to matrix	1.2	1.044
		3	High impact according to matrix	1.4	1.088
		4	Very high impact according to matrix	1.5	1.110
Type/mix of trains	0.09	1	One type of train (e.g. goods train) with similar priority	1	1.000
		2	One type of train with varying priority (e.g. goods trains – hazmat/other materials)	1.05	1.015
		3	More than one type of train with varying priority (e.g. goods trains, passenger trains – excluding Metro trains, hazmat train)	1.15	1.045
		4	More than one type of train with varying priority (e.g. goods trains, passenger trains – including Metro trains, hazmat train)	1.3	1.090
Presence of locomotive depots	0.1	1	No locomotive depot	1	1.000
		2	One locomotive depot (at start or end)	1.1	1.050
		3	Locomotive depot at start and end	1.15	1.075
		4	More than two locomotive depots (including depot(s) en route)	1.2	1.100
Presence of shunting yards/ activities	0.14	1	No shunting en route or if shunting yard or private siding is at end or start of section.	1	1.000
		2	Shunting yards en route	1.05	1.023
		3	Private sidings and/or shunting activities en route	1.25	1.117
		4	Private sidings and/or shunting activities plus shunting yards en route	1.3	1.140
Topography	0.04	1	Karoo landscape; small towns few and far between	1	1.000
		2	Flat terrain; high-density built-up environment	1.05	1.010
		3	Hilly terrain; small towns few and far between	1.1	1.020
		4	Hilly terrain; high-density built-up environment	1.2	1.040

A special procedure had to be adopted to describe the scale points for the moderator, namely “Number of authorisations per shift vs. the number of crossing places”. A work group of experienced TCOs drew up a matrix by considering a wide range of combinations of the number of crossings controlled by a particular centre and the number of authorisations required per shift (See Table 3.8). The legend at the bottom of the figure enables the user to allocate any combination of the two variables to a scale point. These scale points were then quantified using the same procedure as before.

Table 3.8: Impact of number of authorisations and crossing places on planning complexity

		NUMBER OF CROSSING PLACES																								
		2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50+
	2	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
	4	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
	6	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟦	🟦	🟦	🟦	🟦	🟦
	8	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦
	10	🟩	🟩	🟩	🟩	🟩	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦
	12	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦
	14	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦
	16	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟨	🟨	🟨	🟨	🟨
	18	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨
	20	🟦	🟦	🟦	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟥	🟥
OF AUTHORISATIONS	22	🟦	🟦	🟦	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟥	🟥
	24	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟥	🟥	🟥	🟥	🟥
	26	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟥	🟥	🟥	🟥	🟥
	28	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	30	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	32	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	34	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	36	🟨	🟨	🟨	🟨	🟨	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	38	🟨	🟨	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	40	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	42	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	44	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	46	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	48	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	50	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	52	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	54	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	56	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	58	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	60	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	62	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	64	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	66	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	68	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	70	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	72	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	74	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	76	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	78	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	80	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	82	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	84	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	86	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	88	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	90	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	92	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	94	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	96	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	98	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	100+	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥

LEGEND:

COLOUR	**IMPACT ON PLANNING COMPLEXITY**	**SCALE POINT**
Green	Very Low	1
Blue	Low	2
Yellow	Medium	3
Red	High	4

3.10 CALCULATION OF THE MWLI

Using the formula and the weights for the task and moderating factors, the MWLI was calculated for all the train control centres where RTO was being used.

The information on the contributing factors as well as the MWLI calculated for all the centres included in the validation study are shown in Figure 6.2.

3.11 CONCLUDING COMMENTS

In conclusion, a distinction between taskload and workload has been made in the literature and the rationale for considering the MWLI a workload index and not a taskload index, is as follows:

- The distinction between taskload and workload is a very fine one and not many references, where the two concepts are distinguished, could be found. Reinach (2001) differentiates between the two in the following manner:
  - Taskload refers to the number of tasks that dispatchers carry out as part of their job.
  - Workload is generally defined as the interaction between the demands of a given task (or set of tasks) and the ability of the individual operators to meet those demands. Workload takes into account the ability of the operators to meet these demands, and this in turn depends on dispatching experience, training, familiarity with the territory, stress, fatigue, and a host of other factors that affect a dispatcher's ability to meet the task demands.
- It is postulated in this study that the above-mentioned distinction lies on a continuum rather than being an absolute distinction. The MWLI does not only consider task factors but takes into consideration factors such as training (years' experience) and fatigue

(shift). The MWLI therefore leans towards the workload end of the continuum. It was deliberately decided not to include personal factors but to include only information that could be objectively collected. Similar to the FRA's requirement for a tool to collect taskload data (Reinach, 2001), a major consideration in the Spoornet project was that the tasks used in the workload assessment tool must be observable, quantifiable, and quick and unobtrusive to collect.

Authorisation weight	15.00
Tel/radio comms weight	5.00

Centre Id	No of data trans-actions (Rimas + ETA/ETD)	No of authos	Weighted no of authos	No of other tel/radio comms	Weighted no of comms	Total no of actions (weighted)	Moderators											Work-load Index
							Shift worked	Experience	Interface complexity	Running times	Crossing places	Platform Location	Authos vs Crossings	Type and Mix	Loco Depots	Shun-ting	Topo-graphy	
							12%	18%	5%	8%	6%	3%	11%	9%	10%	14%	4%	100%
1	5	34	510	204	1020	1535	1.080	1.030	1.017	1.060	1.000	1.030	1.110	1.045	1.050	1.117	1.020	2629
2	0	9	135	50	250	385	1.080	1.030	1.033	1.080	1.000	1.000	1.000	1.015	1.000	1.117	1.020	553
3	16	97	1455	345	1725	3196	1.080	1.060	1.050	1.030	1.000	1.030	1.110	1.045	1.075	1.117	1.020	5789
4	16	15	225	100	500	741	1.080	1.000	1.050	1.080	1.000	1.000	1.044	1.000	1.050	1.117	1.040	1155
5	0	9	135	9	45	180	1.080	1.030	1.050	1.060	1.000	1.030	1.044	1.000	1.050	1.117	1.000	281
6	4	38	570	271	1355	1929	1.080	1.000	1.050	1.080	1.000	1.012	1.110	1.045	1.050	1.140	1.000	3320
7	3	20	300	130	650	953	1.080	1.030	1.050	1.000	1.000	1.000	1.088	1.015	1.075	1.117	1.000	1476
8	0	35	525	146	730	1255	1.080	1.030	1.033	1.080	1.000	1.030	1.088	1.045	1.050	1.140	1.020	2228
9	24	39	585	297	1485	2094	1.080	1.030	1.050	1.080	1.000	1.012	1.110	1.045	1.075	1.117	1.000	3722
10	18	19	285	366	1830	2133	1.080	1.030	1.050	1.060	1.000	1.000	1.088	1.045	1.050	1.140	1.020	3666
11	7	45	675	325	1625	2307	1.080	1.180	1.050	1.060	1.060	1.030	1.110	1.045	1.075	1.117	1.040	5174
12	5	13	195	264	1320	1520	1.080	1.000	1.000	1.080	1.000	1.000	1.044	1.000	1.050	1.117	1.020	2214
13	0	6	90	8	40	130	1.080	1.000	1.033	1.080	1.000	1.000	1.000	1.000	1.050	1.000	1.040	171
14	2	46	690	222	1110	1802	1.080	1.030	1.050	1.080	1.060	1.030	1.110	1.045	1.050	1.117	1.020	3443
15	0	71	1065	82	410	1475	1.080	1.030	1.033	1.080	1.060	1.030	1.110	1.045	1.050	1.117	1.040	2828
16	0	42	630	327	1635	2265	1.080	1.180	1.017	1.080	1.060	1.030	1.110	1.045	1.050	1.023	1.020	4399
17	10	8	120	23	115	245	1.080	1.060	1.033	1.080	1.000	1.000	1.044	1.045	1.050	1.023	1.020	374
18	20	31	465	261	1305	1790	1.080	1.030	1.017	1.060	1.000	1.012	1.088	1.045	1.075	1.140	1.020	3086
19	9	8	120	88	440	569	1.080	1.060	1.000	1.080	1.000	1.030	1.044	1.045	1.050	1.023	1.020	866
20	0	8	120	91	455	575	1.080	1.030	1.033	1.030	1.000	1.000	1.000	1.090	1.000	1.140	1.020	863

Table 3.9: Calculated MWLI

Table 3.10: Values and interpretation of MWLI elements

	Low centre		*High* centre	
	Value	**Interpretation**	**Value**	**Interpretation**
Number of data transactions	0	Zero data transactions	16	16 data transactions
Number of authorisations	6	6 authorisations	97	97 authorisations
Weighted number of authorisations	90	6 × 15 (weight of authorisation)	1455	97 × 15 (weight of authorisation)
Number of other telephone/radio communications	8	8 communications	345	345 communications
Weighted number of communications	40	8 × 5 (weight of communication)	1725	345 × 5 (weight of communication)
Total number of actions (weighted)	130	∑ weighted authorisations and weighted communications	3196	∑ weighted authorisations and weighted communications
Shift worked	1.080	06:00-14:00 shift	1.080	06:00-18:00 shift
Experience	1.000	7-12 years	1.060	More than 12 years
Interface complexity	1.033	Single interface	1.050	Multi-interface
Running times	1.080	Uniform running times; short	1.030	Varying running times; long
Crossing places	1.000	Uniform crossing places	1.000	Uniform crossing places
Platform location	1.000	No platforms	1.030	More than 50% of crossing places have platforms
Authorisations vs. crossing places	1.000	Low impact according to matrix	1.110	Very high impact according to matrix
Type and mix of trains	1.000	One type of train; similar priority	1.045	More than one type of train with varying priority
Locomotive depots	1.050	One locomotive depot	1.075	Locomotive depot at start and end
Shunting	1.000	No shunting activities	1.117	Private sidings and other shunting
Topography	1.040	Hilly terrain; high-density built-up environment	1.020	Hilly terrain; small towns few and far between
MWLI	**171**		**5789**	

CHAPTER 4

RESULTS

In this chapter the results of the verification study is discussed. Key results are referred to but the detail results are contained in the report attached as Appendix D. Furthermore a sensitivity analysis is performed to provide an indication of the relative contribution to the MWLI of each factors considered.

4.1 CORRELATING THE MENTAL WORKLOAD INDEX WITH PSYCHOPHYSIOLOGICAL PARAMETERS

A requirement was set that the mental workload assessment tool should be a valid assessment of mental workload. This requirement was set to ensure that a mistake made before by Spoornet was not repeated. It was previously decided to use a formula for workload assessment and through prolonged use it was assumed that this formula represented workload, only to discover much later that no proof could be found of any verification that the formula was indeed an indication of workload.

As mentioned earlier, it was initially decided to correlate the NASA-TLX with the MWLI in order to verify the MWLI as a valid assessment of mental workload. Permission to do this was obtained from the developers of the NASA-TLX. The practical circumstances, with varying shift patterns at the different train control centres, meant that it would intrude on the tasks of the TCOs and the notion was therefore rejected.

It was decided that physiological and other data would be collected over an eight-hour period, irrespective of the shift length, starting at 06:00 and finishing at 14:00. For some TCOs this would entail an entire shift but for those who worked a 12-hour shift, this would not be the case.

The pilot study (see Appendix C) showed promising results and, upon approval from the Ethics Committee, the Department of Physiology, University of Pretoria, assisted with the physiological measurements. The author was responsible for the design of the research parameters and approach, setting the requirements of the protocols to be used, as well as the criteria for testing and the desired outcomes.

There was considerable difficulty in finding a suitable instrument that would measure heart-rate variability under the conditions of electromagnetic interference, as was experienced in the pilot study. An instrument had to be imported from the USA, and only after a delay of several months could the field study commence.

Data collection occurred over a period of 9 weeks at 20 train control centres. An initial calculation of ideal sample size indicated that a minimum of 50 centers had to be included in the study to be clinically significant. A maximum of 36 RTO centers existed. All 36 centers were not included due to resource constraints and control over the sample size was therefore problematic. A sample of 20 was considered feasible considering the logistical and resource restrictions. Due to the small sample size, this was not a proper validation study, but rather a pilot verification study through which an approach was developed for a future validation study.

The results of the field study are contained in a report *Validation of the Spoornet Mental Workload Index against an Allostatic Load Index* (see Appendix D for the full report).

4.2 THE VERIFICATION STUDY

The MWL methodology was developed but it still remained to be proven that the index was a valid indicator of mental workload. The purpose of the verification study was to use an independent, objective measure of stress associated with workload, apply it to the different classifications of train control centres, and to determine whether there was supporting evidence for the original postulate.

The approach in the verification study was to examine the physiological changes in TCOs, if any, as indicators of workload stress at the selected train control centres.

The results of these physiological measures were then correlated with the Mental Workload Index to determine whether the calculated index correlated with measures of high and low stress.

4.2.1 Identification of the Experimental Group

Twenty train control centres across South Africa were selected as test sites for the field study.

Using the newly developed MWLI, the train control centres were clustered into high, medium, and low workload by the project team. Using a stratified selection approach, nine *high* and ten *low* workload centres were identified for inclusion in the field study. The scatter gram (Figure 4.1) gives an indication of the clusters that were formed. (A 20th train control centre was included to form part of the field study, although it does not have a pure RTO system – a hybrid system of RTO and track warrant is being used here. This centre was included at the request of the zone operational manager because a decision was pending on whether to split the section and allocate two TCOs instead of one to the section, and objective information was required to make the decision. This train control centre is considered to be a high workload centre.)

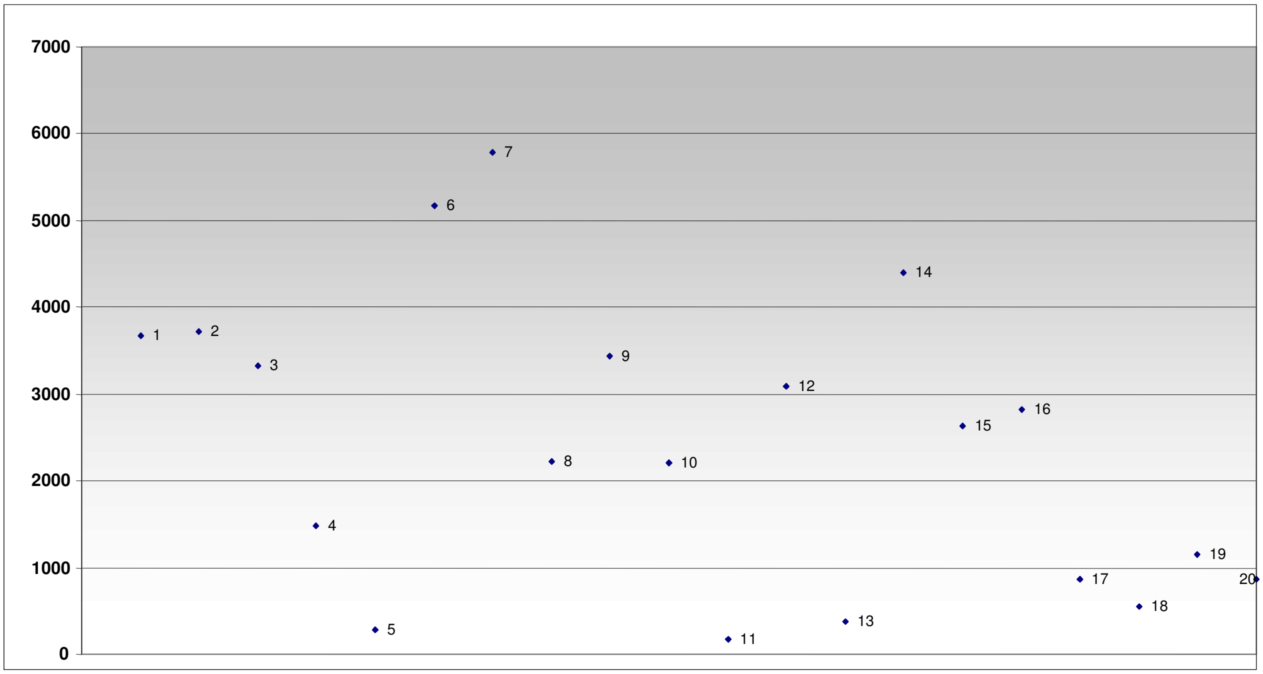


Figure 4.1 Scatter gram of centres used to correlate with MWLI

In the selection of the centres, it was decided to concentrate on distinctive high- and low-classified centres according to the calculated index. Borderline cases were excluded because it was felt that they might not assist in the verification study. Time and cost constraints were also limiting factors. The choice of centres was also influenced by geographical proximity to each other and ease of accessibility, i.e., their proximity to airports.

Table 4.1 contains the train control centres that were selected for inclusion in the verification study and their respective calculated mental workload indexes. To ensure confidentiality the names of the train control centres were omitted and identification numbers were allocated.

Table 4.1: Train control centres included in the verification study

ID	**Name of centre/section** (Deleted to maintain confidentiality	**Calculated MWL Index**	**Rating**
1		2629	H
2		553	L
3		5789	H
4		1155	L
5		281	L
6		3320	H
7		1476	L
8		2228	L
9		3722	H
10		3666	H
11		5174	H
12		2214	L
13		171	L
14		3443	H
15		2828	H

ID	Name of centre/section (Deleted to maintain confidentiality	Calculated MWL Index	Rating
16		4399	H
17		374	L
18		3086	L
19		866	L
20		863	L

4.3 MEASUREMENT TECHNIQUES AND METHODOLOGY

4.3.1 Measurement Techniques and Instruments

The following measurements were performed during the field study:

- **Cohen's Measure of Perceived Stress** – Subjects completed the stress questionnaire, with the assistance of the field worker, to assess their non-specific appraised stress during the last month. The questionnaire consists of 14 items of which seven are positively formulated and seven are negatively formulated. (See Appendix A.)
- **Blood pressure** – Diastolic (DBP) and systolic (SBP) blood pressure were measured with a digital electronic blood pressure meter (ALP K2, model DS-125D, Japan). Three consecutive measurements were taken, the first at the commencement of the shift and every two hours thereafter during the shift. The average of the three measurements is reported.
- **Heart Rate Recording and Analysis** – A Direct Wired POLAR ® heart rate belt (Mini Mitter Co. Inc. Bend, OR, USA) was attached to the subject. Heart rate data was transferred instantaneously via a direct wire to the Mini-Logger® Series 2000 (Mini Mitter Co. Inc.) and stored until download. Heart rate data, recorded continuously during the eight-hour exposure, was then downloaded to a laptop computer using the Mini-Log™ 2000W for Windows® (Mini Mitter Co. Inc. Bend, OR, USA).
- **Cortisol Collection and Data Analysis** – Cortisol is found in the blood or saliva. Free salivary cortisol was determined with a Salivary Cortisol ELISA kit (DX-SLV-2930, AEC Amersham Pty, Ltd., South Africa). Saliva was collected in a clean collection test tube at

the start of the shift and every two hours thereafter until 14:00 and then stored on ice to centrifuge later.

- **BMI** was calculated for all subjects.
- **TLA** – A real-time TLA was recorded during the same shift that the physiological measurements were performed. (See Appendix B for TLA template.)

The same observer was responsible for the TLA and the physiological measurements of all the subjects tested at all 20 train control centres. This was done to ensure objectivity and to eliminate any contamination of data that may have resulted from the different approaches of different individuals.

4.3.2 Measurement Methodology

Testing commenced at the start of the shift at 06:00. In four of the cases the TCOs and area management had an agreement that the shift would commence only at 07:00. In these instances testing commenced at 07:00. The reasons for deciding on this time period were:

- Physiological measures had to be taken during the day shift because of the body's physiological changes, including changes in cortisol levels, during the night due to the circadian rhythm.
- A period of eight hours was chosen due to the fact that in most of the train control centres eight-hour shifts were worked. An eight-hour period was therefore used consistently even if a 12-hour shift was worked.

All operations managers were informed of the study a month prior to commencement of the field study and were requested to inform TCOs of the study and its aims. Employee representative organisations were involved throughout the development process and ensured communication with and cooperation from their members.

Upon commencing their shift, the subjects performed their handing over duties and were then informed by the observer about the exact nature of the physiological measurements. The experimental procedures and aims were explained, consent to participate was obtained where after testing commenced. Participation in the study was voluntary and TCOs could choose not to be tested. There were no instances of refusal to participate. Although the field workers knew that the workload at the different train control centres differed, i.e., nine were high and ten were low workload centres, they did not have insight into which centres were high workload and which were low workload centres until the completion of the statistical analyses of the results.

4.3.3 Measurement Administration

The test protocol was submitted to and approved by the Ethics Committee, Faculty of Medicine, University of Pretoria and Pretoria Academic Hospitals (Number IPA-41 - 150/2000). Before testing commenced, all participants in the validation study were informed of the measurements by their managers. On the day of the study itself they were again informed of these by the field worker.

The tests were conducted by a postgraduate student in Physiology from the University of Pretoria under the supervision of a senior member of that department.

4.4 DATA ANALYSIS APPROACH

The statistics used in the analysis of the results can be summarised as follows:

All values are represented as mean ± standard deviation. A Two Sample T-test was used to determine whether differences existed in terms of allostatic load variables between the identified and high stress centres respectively. A two-way repeated measures Analysis of Variance (AOV) was used to evaluate between-subject differences (low and high stress centres) and within-subject differences (responses over an eight-hour shift). Group*Time interaction was also evaluated to determine whether both groups responded in a similar fashion over the eight-hour shift. If between- and within-subject analyses did not indicate any significant differences, but Group*Time interaction did, it would have meant that the two groups responded differently to the stressor over time.

All data and statistical analyses are contained in a separate report (Appendix D). The Statistix Ver. 8 program was used for the statistical analyses. The overall approach that was followed, i.e. the collection of the raw data, measurements, data and statistical analyses, conformed to the set requirements and provided the necessary objective information.

4.5 RESULTS OF THE VERIFICATION STUDY

Twenty train control officers at various train control centres were selected to take part in the physiological verification of stress levels manifested as a result of mental load. Selection was performed on the basis of an estimation of the stressor impact at the different centres as calculated by the MWLI. The brief for the verification study was to investigate whether the MWLI was supported by the values of physiological stress indicators.

All raw data, results of statistical analyses, findings, and discussion concerning these are contained in the comprehensive report attached as Appendix D.

The mean physiological values over the shifts for sub-groupings according to the MWLI were compared to see whether physiological differences could be found between individuals that had been grouped according to high and low MWLI values. The parameters that were included in this comparison comprised factors that could either reflect the stress reactivity in individuals, or the wear and tear as a result of chronic exposure to stressors, or factors that could influence the aforementioned two types of indicators. Comparisons were made for age, mass, height, body mass index (BMI), surface area (SA), systolic blood pressure, diastolic blood pressure, mean arterial pressure, pulse pressure, heart-rate variability variables, smoking, length of previous shift, length of test shift, number of years experience at a particular station, shift preferences and timeline analyses of the shifts when the physiological recordings were made.

Very few parameters showed any significant difference between the high- and low-stress groups. Only the TLA, which is built into the MWLI, showed, as could be expected, statistical differences between the high- and low-workload groups throughout. In short, the values of the physiological stress parameters did not mirror the MWLI. To see whether exposure to high workloads caused higher anticipation stress in workers, the arrival values (before the workload could have had any effect) between the two groups were compared for those parameters that were measured on arrival. Again, no significant difference could be found between high- and low-stress groups.

Statistical power analysis was not conducted due to the small sample size.

In order to ensure a sharp distinction between high and low workloads, as calculated by the MWLI, the values of only those individuals at the six highest and six lowest ranked train control centres were compared. The statistical analyses of this data once again showed that physiological stress values did not mirror the activity intensity as described by the MWLI. Very significant correlations were again seen between timeline analyses and experimental values. These correlations between the TLA factors and the early-morning values were probably only a reflection of the fact that timeline analyses were built into the MWLI. This nevertheless confirmed the fact that the MWLI is a good reflection of activities at the specific centres.

The next step was to look at the reactivity over the duration of shifts. When the high and low MWLI groups were compared in terms of the way they physiologically reacted to the workload over the total shifts, hardly any correlations were seen between the workload and reactivity. Many factors probably contributed to the fact that the results of the MWLI were not mirrored by the stress levels of the workers. The most likely contributing factors were probably individual differences such as age, health status, years of experience, and gender. The populations at the various stations differed significantly with regard to such aspects, and with the small experimental group size of this project, such differences could have nullified any significant differences. An additional confounding factor was the fact that three of the twenty workers were female. There are indications that females are perhaps more stress-responsive than males. This, however, is contradictory as there are certain factors that protect females against the negative effects of high-stress system activation. In this study the females had lower stress levels but were also younger than the average TCO – a fact that could very well explain their low allostatic loads.

There are several other reasons why the differing workloads of stations might not necessarily have been reflected in the physiological values of workers and in the changes in response to workload at the specific stations. The role of the adrenal medulla, an extension of the sympathetic nervous system, needs to be mentioned here. The secretory response of the adrenal medulla to stressors is dependent on the previous stress history of the individual. For instance, with exposure to the same stressor each day over an extended period of time, subsequent exposure to the same stressor may induce a lower adrenal medullary response. This response is known as habituation. However, if the same individual who has been stressed by a particular stressor over an extended period of time is suddenly, in addition to the previous stressor, exposed to yet another, novel, stressor the adrenal medulla secretory response may be significantly larger. This response is known as sensitisation.

Habituation to known stressors may occur as a result of repeated exposure to the same stressor, and TCOs may therefore adjust to high workloads if they are subjected to them over extended periods of time. It is only when novel stressors are introduced into the work environment, such as dramatic changes to the normal routine, that sensitisation and the effects of high allostatic loads, with their concomitant repercussions for performance and mental and physical health, might be noticed. It is self-explanatory that changes in Spoornet infrastructure and policies may result in such novel stressors.

It would seem that acute stress superimposed on chronic stress is dependent on the individual’s familiarity with the type of stressor. It appears, then, that the response to an acute stressor in the chronically stressed individual would be less than expected if the stressor is homotypic (a repetition of what caused the chronic stress response). In contrast, if a heterotypic stressor were applied to a chronically stressed individual, the response would be greater. It is possible, however, that cross-tolerance to stressors may develop, especially with those stressors that involve the same neurological pathways. In theory, this can be extrapolated to the working situation where high cortisol levels might perhaps not increase significantly as a result of increased levels of homotypic stressors, such as increases in the amount of work to which the individual is accustomed.

In contrast, cortisol levels may increase when atypical stressors are encountered. The implication that can be deduced from this (i.e., that the amplitude of the stressor-induced increase in cortisol levels and the circadian amplitude difference to additional homotypic stressors may be lower during chronic high stress) is that individuals with chronically high cortisol levels will not show the expected increase from basal values when stress levels increase and that this could be ascribed to the values being higher than normal. To extrapolate this to TCOs, one could expect that habituation to their respective workloads could rule out the development of significant stress differences between the different workload groups. The magnitude of the stress reactions from baseline to stimulated responses could be negated by above normal baseline values.

The next step was to identify high and low stress groups on the grounds of physiological variables. A comparison was made between the values of the MWLI at the various centres on the one hand, and the physiological indicators of stress in the workers at the different stations, on the other. The possibility of separating the workers at the different centres into high- and low-stress groups on the basis of the values of their physiological parameters was investigated. The analyses, as previously mentioned, were performed in terms of an adapted allostatic load measurement where the data used in the calculation of the numerical value for each individual factor or parameter of the allostatic load were derived from the mean of the values obtained over the work shift. The reason for this was that baseline values in this group could not really be seen as baseline because all of them appeared to arrive with values higher than those recorded during the experimental procedure. Many factors could have contributed to this observation, most noticeably the fact that their arrival time coincided with higher cortisol levels relating to the circadian rhythm. Subsequent measurements were performed at times when the circadian values were already lower. The anticipation of stress before commencing a shift could have been a further contributing factor. In addition, it is known that individuals in high-stress jobs seldom recover to baseline values during their time off.

There can be no doubt that the stressors in the case of TCOs should be considered uncontrollable as the number of trains and other activities are predetermined by the job at hand and are not under their control. The fact that the stressor, in this case, is not a once off stressor but that TCOs live with these stressors for the major part of every day supported the decision to use the mean value over the shift worked.

As previously mentioned, in the process of subdividing TCO workers into high- and low-stress groups based on physiological values, the values of the following parameters were included: salivary cortisol and BMI as indices of hypothalamo-pituitary-adrenocortical axis activity; systolic and diastolic pressure as a measure of cardiovascular activity which largely reflects sympatho-adrenomedullary axis activation; heart rate as an indicator of sympatho-adrenomedullary activation; and heart-rate variability as indicators of autonomic activity. The values for each individual for each of the five indicators were classified according to the 50th percentile. Allostatic load was calculated by summing up the number of parameters for which the subjects fell into the highest risk, i.e., above the 50th percentile. Subjects were subsequently ranked according to their total hits.

Three models, Model A (including the values of cortisol, systolic blood pressure, diastolic blood pressure, heart rate, low frequency and BMI), Model B (cortisol, systolic, diastolic, heart rate and BMI) and Model C (cortisol, systolic, diastolic and heart rate) were developed and tested, and the TCOs were subdivided into high- or low-stress groups. When the TCOs were subdivided into high- and low-stress groups and the differences for all physiological parameters (not only those included in the three models) were tested for significant differences between the groups of individuals subdivided into the high- and low–stress groups according to Models A, B and C, it was seen that Model B was superior to Models A and C. This was supported by the fact that for Model A, 10 out of 17, for Model C, 8 out of 17, and for Model B, 12 out of 17 physiological variables differed significantly ($p < 0.05$) between the high-stress groups and the low-stress groups (two sample T-tests). The significant differences were also generally higher for Model B than for the other two.

In the final analysis, subdivisions of workstations on the basis of the MWLI were compared to subdivisions of the individuals at those stations and to perceptions of the observer regarding the stressor value of the workstation. A 90% agreement was found between the MWLI and observer perception (this was a subjective perception and not based on the sum of the activity as determined by timeline analyses) and a 60% agreement was found between the subdivisions into high- and low-stress stations by MWLI, on the one hand, and subdivisions based on Model B, on the other.

TLA is built into the MWLI and one would, therefore, expect to find some kind of correlation between TLA and MWLI. In an attempt to investigate the strength of the correlation and to see whether the MWLI has a significant advantage over simple TLA, correlations were tested between, on the one hand, the MWLI and, on the other hand, the individual factors in TLA as well as two combinations of TLA factors. The correlations between the MWLI and the two combinations of factors were $r = 0.7986$; $p = 0.0001$ (trains and authorisations) and $r = 0.9110$; $p = 0.0001$ (radio, telephone, schedules), and all individual factors also showed significant correlations. It would appear, therefore, that the TLA gives very much the same information as that derived from calculating the model. TLA as a stand-alone technique would, however, not be acceptable as a valid and reliable tool unless it had been validated through a scientific study. Furthermore, TLA on its own does not provide insight into the reason(s) for possible mental overload.

4.6 SENSITIVITY ANALYSIS AND OPTIMISATION OF THE MWLI

4.6.1 Sensitivity Analysis

Sensitivity Analysis is used to determine how sensitive the results of a study or systematic review are to changes in the manner in which a study was conducted. Sensitivity analyses are used to assess how robust the results are in the face of uncertain decisions or assumptions about the data and the methods that were used (Saltelli, Chan, and Scott, 2000).

Sensitivity analysis is also used to determine how ‘sensitive’ a model is to changes in the value of the parameters of the model and to changes in the structure of the model. Sensitivity analysis helps to build confidence in the model by studying the uncertainties that are often associated with parameters in models. Many parameters in system dynamics models represent quantities that, in the real world, are very difficult, or even impossible to measure with a great deal of accuracy. Therefore, when building a system dynamics model, modellers are usually slightly uncertain about the parameter values they choose and must therefore use estimates. Sensitivity analysis allows them to determine what level of accuracy is necessary for a parameter to make the model sufficiently useful and valid. If the tests reveal that the model is insensitive, then it may be possible to use an estimate rather than a value and achieve greater precision. Sensitivity analysis can also indicate which parameter values are reasonable to use in the model (Breierova and Choudhari, 1996).

The objective of sensitivity analysis here is to provide an indication of the relative contribution to the final MWLI of each of the factors considered. The sensitivity of the MWLI to changes in the *task factor* and *moderating factor* weights was determined by a first-order sensitivity analysis, using the measured data and changing only one parameter at a time while keeping all the other parameters at their nominal values. Sensitivity in this context is expressed as a ratio of the difference between the two factors. The nominal weights, relative to which the changes were considered, were taken as the weights determined by the consultative process described earlier. Table 4.2 shows the nominal weights of each of the task factors and moderating factors.

Table 4.2: Nominal weights of task and moderating factors

Task Factors	**Symbol**	**Nominal weight**
1. Data transactions	D	1
2. Communications	C	5
3. Authorisations	A	15
Moderating Factors		**Nominal weight %**
1. Shift type	M1	12%
2. Experience as a TCO on the particular system	M2	18%
3. Planning complexity		Total 33%
3.1 Interface complexity	M3	5%
3.2 Running times between crossing places	M4	8%
3.3 Types of crossing places	M5	6%
3.4 Locations of platforms	M6	3%
3.5 Number of authorisations per shift vs. number of crossing	M7	11%

places		
4. Inherent difficulty of the section		Total 37%
4.1 Type/mix of trains	M8	9%
4.2 Presence of locomotive depots	M9	10%
4.3 Presence of shunting yards/activities	M10	14%
4.4 Topography	M11	4%

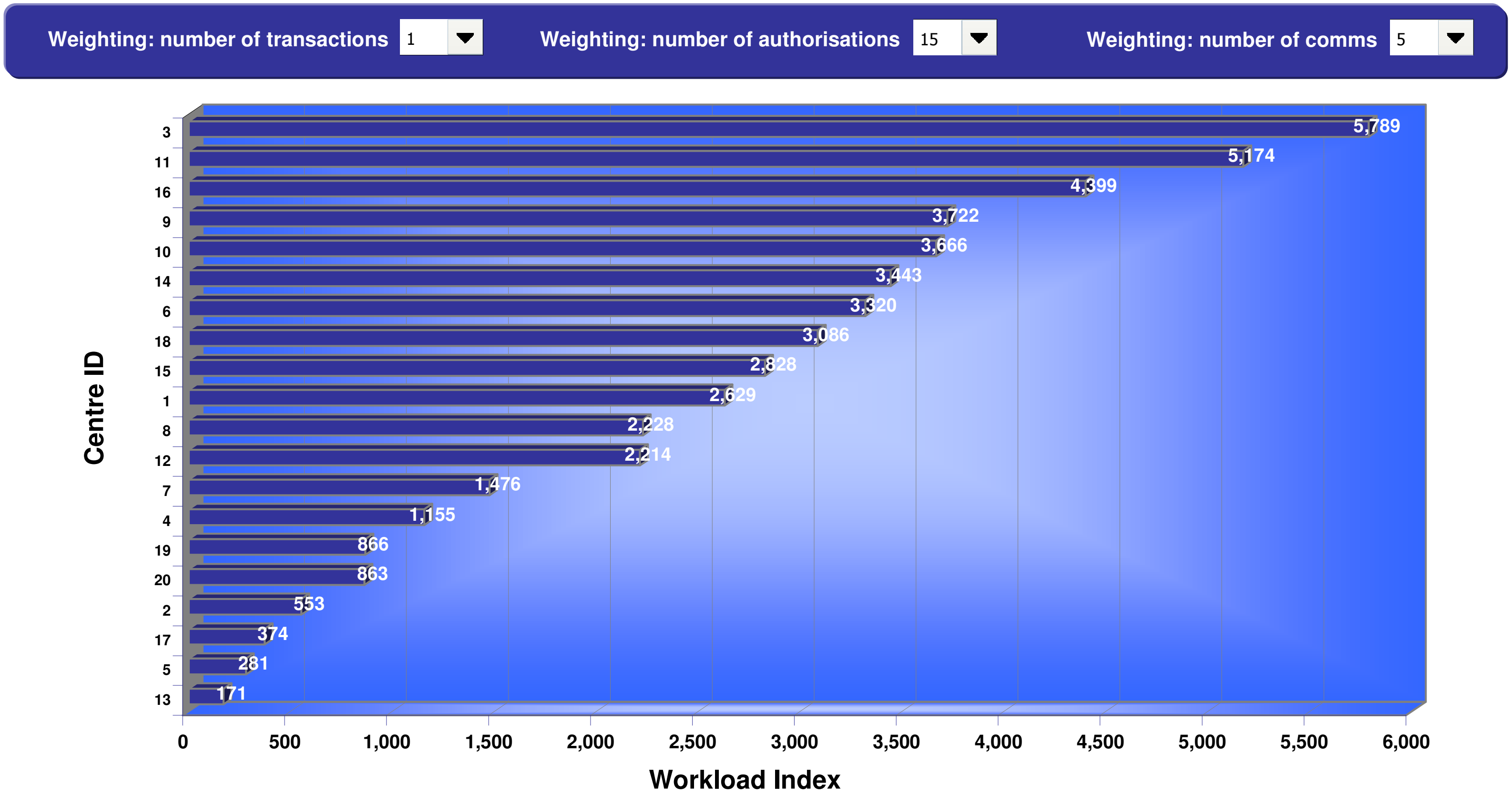
Sensitivity of the MWLI was defined as

$$S_{PQ} = \frac{\Delta P / P_0}{\Delta Q / Q_0}, \qquad \text{(Generic Formula)},$$

with $\Delta P / P_0$ indicating the change in the MWLI in a particular group of train centres, relative to the nominal value P_0 of the MWLI for the same group, and with $\Delta Q / Q_0$ indicating the change in the weight of a particular parameter, relative to the nominal weight Q_0 of the same factor. The output variable P can be P=H, M or L representing the mean of the MWLI for the high (H), medium (M) and low (L) groups of train control centres, while the input variable Q can be index Q=A, C or D representing the task factors A=Authorisations, C=Communications, D=Data transactions or Q=M1…….M11 representing the moderating factors.

The train control centres were grouped as follows according to the calculated MWLI for the nominal values of the task and moderating factors. (See Figure 4.2 for H, M, L groupings):

High:	$\text{MWLI} > 3000$
Medium:	$1000 < \text{MWLI} \leq 3000$
Low:	$\text{MWLI} < 1000$



Weighting: number of transactions 1

Weighting: number of authorisations 15

Weighting: number of comms 5

Figure 4.2: High, medium and low groupings for calculated MWLI

The mean values of the MWLI for each of the groups (H, M, L) for the nominal values are given in Table 4.3.

Table 4.3: Mean value of MWLI for high, medium and low groups

Group	**Mean MWLI**	**Centre ID**
High	4075	18, 6, 4, 10, 9, 16, 11, 3
Medium	2088	4, 7, 12, 8, 1, 15
Low	518	13, 5, 17, 2, 20, 19

The results of the sensitivity analysis for the above procedure are given in Table 4.4.

Table 4.4: Results of sensitivity analysis of task and moderating factors for H, M, L groups

				Group		
Task Factor	Nominal value	Δ-Value		High	Medium	Low
A	15	16		0.31	0.39	0.34
C	5	6		0.69	0.60	0.65
D	1	2		0.50	0.36	0.93
Moderating Factor	Nominal value	Δ-Value		High	Medium	Low
M1	12%	13%		0.07	0.07	0.07
M2	18%	19%		0.07	0.02	0.04
M3	5%	6%		0.04	0.03	0.02
M4	8%	9%		0.06	0.06	0.06
M5	6%	7%		0.02	0.01	0
M6	3%	4%		0.02	0.02	0.01
M7	11%	12%		0.10	0.08	0.02
M8	9%	10%		0.04	0.03	0.04
M9	10%	11%		0.06	0.05	0.03
M10	14%	15%		0.10	0.11	0.07
M11	4%	5%		0.02	0.02	0.02

The sensitivities are expressed as fractions of 1 with S=0 indicating no change in MWLI for a change in the selected weight, S=1 indicating a 1:1 change in the MWLI for a change in the selected weight, and S=0.5 indicating that the MWLI changes by 50% from the nominal value for a 100% change in the selected weight.

In each case the weight was varied by one unit while all the other weights were kept at their nominal values. It should be noted that each of the calculated sensitivities is to some extent dependent on the nominal values of the other parameters. The results of a multivariate sensitivity analysis, in which more than one parameter is changed at a time, is more difficult to interpret and is considered beyond the scope of this study.

The moderating factors weights could not be analysed in the same way as the task factor weights, since there was a unique set of 11 nominal values for each train control centre. In the case of moderating factors the maximum contribution of each factor to the MWLI was considered as the nominal weight value.

The first-order sensitivity analysis revealed the following trends and their possible interpretations:

Task Factors:

- The sensitivity of the MWLI index in terms of the task factor with the highest weighted contribution, i.e. authorisations, with a nominal weight of 15%, has the lowest sensitivity with respect to all three the task factors. Further, the sensitivity is about the same (~0.3) for all three groups, high, medium and low. This indicates that the MWLI is least affected by a change in the number of authorisations, despite the fact that in the opinions of the expert advisory group involved in the consensus-driven design of the MWLI this task factor was considered to make the largest contribution to the MWLI. It appears, based on the results of the sensitivity analysis, that the perceptions, on which the relative contributions of the task factors to the MWLI were based, might have been incorrect.
- The MWLI in all three groups has a sensitivity to changes in the communications task factor of 0.6, which is roughly double that of the authorisations task factor, matching the intuitive observation that an incremental increase in the number of communications has

a greater effect on the workload of TCOs in all groups than an increase in the number of authorisations

- The sensitivity of the MWLI to changes in the number of data transactions shows the greatest variation (0.3 – 0.9). It is interesting to note that this factor shows the highest sensitivity in the low MWLI group (0.9). The interpretation of this could be that in a low workload train control centre a change in only one unit in the data transactions task factor results in a change of more than 90% in the MWLI. This fits the intuitive observation that an increase in the data transactions significantly increases the workload of TCOs in low workload centres.

Moderating Factors:

- The moderating factors with the lowest percentage contribution to MWLI (<8%) showed consistency in that the MWLI had the least sensitivity (<0.04) across all three groups. This was to be expected as these factors were considered to be the least important contributors to MWLI.
- The sensitivity of the MWLI to the moderating factor *running time between crossing places* (M4), which had a nominal 8% contribution, was higher than it was to the moderating factor *type/mix of trains* (M8) which had a nominal 9% contribution. The variation in the sensitivity was about 0.02 but there was consistency across all three groups.
- The moderating factors with the greatest variation in sensitivity across the three groups were also the factors with a higher percentage contribution (M2=18%; M7=11%; M9=10%; M10=14%). These factors also showed a higher MWLI sensitivity for the high workload group, except for M10, (*Presence of shunting yards/activities*) which showed a higher MWLI sensitivity for the medium workload group.
- The only moderating factor that presented an exception to the above trend was factor M1 (*shift type*) with a weight of 12%, which showed a lower comparative MWLI sensitivity (0.07) and no variation across the three groups.
- Except for shift type, it therefore seems that the moderating factors with the highest contribution to MWLI also showed the highest sensitivity. This can be taken to confirm the nominal contributions of the moderating factors being allocated consistently by the expert advisory group.

4.6.2 Optimisation of the MWLI

The MWLI is a multiparameter index that has no well defined minimum or maximum. Hence the optimum cannot be mathematically and quantitatively determined from its definition. The optimisation of the MWLI can only be addressed in a qualitative sense in terms of it being most appropriate for any given set of circumstances. The optimisation of the MWLI is beyond the scope of this study, inasmuch as it needs the long-term extensive use of the MWLI and an evaluation of its appropriateness needs to be made in terms of indirect and subjective indicators of mental workload – such as operator experience, turnover of staff, accident frequency, long-term health effects, sick leave and absenteeism – being correlated with the high values of the MWLI as determined by the proposed index.

4.7 SUMMARY OF RESULTS

The workloads at the various stations, as predicted by the MWLI, were not reflected by either the adapted allostatic load (including all measured parameters of the individuals working at those stations) or by changes in stress levels over shifts. However, in developing the three models, consisting of different combinations of allostatic load indicators, there was a 60% correspondence between the workstations subdivided into low and high workloads, according to the MWLI, and the subdivision of TCOs at the corresponding train control centres into high and low stress according to Model B. Model B was based on the mean of the values taken over the duration of the shift for cortisol, systolic pressure, diastolic pressure, heart rate and BMI. It can thus be said that the combination of physiological parameters used in Model B supports the potential usefulness of the MWLI to differentiate between high and low stress due to mental workload.

CHAPTER 5

DISCUSSION

In this chapter the hypothesis and research questions are critically reviewed in terms of the results and the literature. In closing, possible applications of the Mental Workload Index are discussed.

5.1 HYPOTHESIS AND RESEARCH QUESTIONS

The limiting factors such as sample size, heterogeneity of the sample and the habituation phenomenon, discussed in the previous chapter, compromised the significance of the results considerably. Despite this promising results were obtained and it can be concluded that this study has, with limited success, proved that it is possible to develop a preliminary objective and non-intrusive method to assess mental workload in a specific environment.

The scientific hypothesis that was tested in this research is:
The objective mental workload index as developed in this thesis can differentiate between high and low imposed workload train control centres.
The hypothesis can be conditionally accepted namely for this particular study with this particular sample, but it has not been proven correct for other circumstances.

The following research questions were posed:

- Which operator tasks could potentially be associated with mental workload?
- Which moderating factors, which would either increase or decrease mental workload should be considered and how can their respective moderating effects be determined?
- Which parameters should be considered to develop a measure of a mental workload that is completely objective and requires no estimation of experienced load by the operator?
- Can the proposed index differentiate between workload levels that could potentially be experienced at different train control centres?
- Which physiological measurements provide an objective assessment of mental workload?

The research questions are critically evaluated in order to determine whether they have been adequately addressed.

- Which operator tasks could potentially be associated with mental workload?
    - This question was answered adequately.
    - Tasks were identified by observing TCOs at work and then developing a TLA template containing these tasks. (See Chapter 3 – 3.7 and 3.8).
    - The tasks were verified by experienced TCOs as a true reflection of their daily activities.

- Which moderating factors, which would either increase or decrease mental workload should be considered and how can their respective moderating effects be determined?
    - This question was addressed satisfactory, but the contribution of the moderating factors to overall mental workload does not seem significant.
    - Through a process of user participation and after several iterations, moderating factors were identified. (See Chapter 3 – 3.7 and 3.8).
    - The contribution of each of these factors to mental workload was then determined using a modified nominal group technique. (See Chapter 3 – 3.9).
    - These values were normalised and scale points allocated to the descriptors. (See Chapter 3 – 3.9).

- Which parameters should be considered to develop a measure of a mental workload that is completely objective and requires no estimation of experienced load by the operator?
    - The answer to this question is not a definite affirmative or negative.
    - A qualified affirmative is probably the best answer based on the results obtained from this sample.
    - The shortcomings have been mentioned and better correlations are possible with proper sampling of subjects. (See Chapter 4 – 4.5).

- Can the proposed index differentiate between workload levels that could potentially be experienced at different train control centres?
    - This question was answered to an extent.

  - With a refining of the groupings of parameters a correlation of 60% was obtained between these parameters and a high and low stress group of TCOs respectively. (See Chapter 4 - 4.5).
  - Proper sampling would probably improve the correlations. This is addressed in Chapter 6 – 6.2.

- Which physiological measurements provide an objective assessment of mental workload?
  - This question could not be answered.
  - Based on the literature it was considered to be a good means of validation Physiological variables change with mental activity. It is therefore seen to be well suited as indicators of mental workload demand (Rasmussen, 1979).
  - Furthermore, because the changes that occur are involuntary and subjects are not requested to execute any extra overt behaviour it seems a suitable means of validation. (See Chapter 2 – 2.8.4).
  - Physiological measurements of stress associated with mental workload is however hard to isolate and because of other operator states that also impact on these parameters will always have a risk of being contaminated.
  - In this study the TCOs are all shiftworkers and therefore are subject to fatigue due to their schedules.
  - They perform safety-critical tasks and one would assume that the levels of stress would be elevated.
  - As explained in Chapter 4 (4.5) the relationship between stress and mental workload is a complex one and in retrospect this relationship is potentially not strong enough or weakened by related variables to such an extent that it should not be used as a basis for validation

It is concluded that there was little evidence of high levels of stress in TCOs, either from subjective stress ratings or from salivary cortisol levels or any of the other physiological parameters. These levels were well within normal levels for adults. These results should not be interpreted as an indication that workplace stress does not exist. It is more probable that data was collected too infrequently in this rapidly changing environment and may not have captured the changing workload and related stress.

It is significant that Popkin et al. (2001) documented results that are comparable to the results found here: 95% of cortisol samples in their study fell below the upper limits of normal cortisol levels. It is possible that these levels were suppressed due to either habituation or burn-out of the cortisol producing system within the brain. Either of these conditions could result if individuals are placed under high levels of chronic stress.

The physiological and self-rating data did not indicate that the dispatchers were under a high stress load. This is surprising given the number of items that were listed as contributing to their at-work stress. Anecdotal evidence, as well as the safety critical nature of the job, would also suggest a high-stress environment. It may be that dispatchers are a self-selected group that have either a higher tolerance to stressors or better coping mechanisms (Popkin et al., 2001).

Significant correlations between TLA and MWLI were found, which was to be expected. This means that the MWLI gives good reflection of the activities at the specific centres. Given these correlations it is possible to utilise only that part of the MWLI that calculates task factors, to determine the potential task load a train control centre.

5.2 CRITICAL EVALUATION OF THE CONTRIBUTION OF THE MENTAL WORKLOAD INDEX

Operator states such as fatigue and stress that are related to mental workload rarely present themselves as a single challenge in research. They are all relevant in environments of operational research (Desmond and Hancock, 2001).

The MWLI is an attempt to obtain a measure of ‘pure’ mental workload factors, but by its nature mental workload is always mediated by factors internal and external to the operator. An attempt was made to address mediating/ moderating factors in the MWLI, but a shortcoming of the index is that no other operator states were added as moderating factors. There was good reason for this namely the aim to develop a tool that operational personnel and others could use to objectively, reliably and quickly obtain taskload data. The effect of this shortcoming however is that the MWLI could potentially only measure operator stress and not mental workload.

The experience of mental workload depends on the individual and is the result of the interaction between the operator and task. Consequently, the same task demands do not result in an equal level of workload for all individuals (De Waard, 1996). A shortcoming of the MWLI is that it potentially underestimates the interactive nature of mental workload and a too simplistic view was enforced due to its emphasis on task and other objective factors. This simplistic view was acknowledged and possible criticism was anticipated (Chapter 2 – 2.2.3). Retrospectively, a more viable approach would have been to develop an index that would provide a calculation of only the task load at a specific train control centre rather than a mental workload index which implies operator states and other human factors.

It is necessary to ask whether either the development of a MWLI or the use of the physiological parameters of the workers at the various centres gave significantly better estimates of the stress levels at the stations than the use of only the TLA. At this stage the answer seems to be negative. Although the MWLI is a good reflection of the activities at the stations, and Model B (in the verification study) supports the use of the index, neither would appear to have an advantage over the TLA, except for the fact that the TLA is considered a subjective technique and therefore does not comply with the criteria of objectivity and scientific validation.

Another question, in view of the high correspondence between the MWLI and the subjective perceptions of an observer, is whether the workload could not be estimated simply through observation, without the use of the TLA, physiological measurements or calculation of the MWLI. Once again, the answer is likely to be negative as estimations of workload can differ, as they may be dependent on the ability, commitment, and objectivity of the observer.

It would seem from the literature and from the results that despite its shortcomings, the MWLI does offer some benefits over existing mental workload measuring techniques that potentially make it a useful index, provided that further research is done on its validity and application in other contexts.

Secondary measures of spare mental capacity offer the benefit of high face validity and a level of standardisation. Some of these measures can be applied to different primary tasks and the workload measure will be given in the same units. It however suffers on the intrusiveness criterion.

The MWLI also has high face validity due to the inclusion of task factors. It satisfies the intrusiveness criterion and information is collected unobtrusively and it can even be used as a desktop exercise.

Subjective rating techniques offer the advantages of not disrupting the task and ease of application. In comparison, the MWLI does not require any input from the operator, is completely non-intrusive and does not have the disadvantage of subjective techniques, namely that it is based on operators' perceptions and possibly does not diagnostically reflect mental workload.

Specific criticism of well-known subjective techniques such as NASA-TLX and SWAT, relate to the fact that these have been simulated in the military environment and is not considered relevant for TCOs as it does not capture all the relevant aspects of their tasks. The language and application is also not considered appropriate for use in real-time in the field with civilian populations. These instruments also assess the workload of a single task whereas TCOs complete multiple tasks concurrently (Pickup et. al., 2005).

The MWLI has distinct advantages over these shortcomings as it addresses the multiple concurrent tasks, is easy to apply and does not utilise subjective information.

The advantages of the MWLI over physiological measurements are evident. Ease of use, cost and the necessity of specialised field personnel make the use of these measurements virtually impossible.

In conclusion, the MWLI can potentially be used as a method of choice, provided that existing shortcomings, especially around its validation, are addressed in further research.

The sensitivity analysis in Chapter 4 was done to determine how sensitive the index is to changes in the value of parameters. It provides an indication of the relative contribution to the MWLI of each of the factors.

The first-order sensitivity analysis provided an independent confirmation of the assumptions about the data and the methods that were used in the design of the MWLI.
The conclusion is that the MWLI, as used in this study, could reliably discriminate between high and low mental workload centres based on traffic volumes. It is therefore a useful indicator and predictor of the stress associated with workload. The calculated MWLI should not be interpreted as an absolute value but rather as an indicator that should prompt managers of the system to take preventative action in order to ensure safe train operations.

A further aim of the study was to develop a methodology that would stand up to the tests of validity and reliability. This entailed that the criteria set by other researchers for mental workload assessment techniques should be met. Based on the results obtained with the MWLI, it is argued that this methodology meets the following criteria:

1) <u>Sensitivity</u>: The MWLI shows sensitivity to changes in task difficulty or resource demand, i.e., an increase in demand on the task elements has a direct and visible impact on the overall workload.

2) <u>Selectivity</u>: The index should be selectively sensitive only to differences in resource demand and not to changes in factors such as physical load or emotional stress, which may be unrelated to mental workload or information-processing ability. At the outset of the study it was made clear that the methodology should exclude personal factors and that these should be identified and addressed through other processes since they might impact on safe operations.

3) <u>Intrusiveness</u>: The MWLI does not interfere with, contaminate or disrupt performance of the primary task. The MWLI can be performed off-site once the required information has been captured. It does not require the direct involvement of the operator.

4) Operator acceptance: It was imperative that the TCOs supported and accepted the methodology and that they believed that the MWLI measures what it claims to measure. The MWLI was developed with the inputs from users of the system and, therefore, has gained credibility and acceptance.

5) Face validity: Because the MWLI uses task-related information and information that relates to operator capability as well as moderating factors that could add to or lessen operator stress and task difficulty, it has face validity.

6) Affordability: Costs were incurred in the development of the MWLI. There is no further cost in the implementation or use of the MWLI.

7) Implementation requirements: The MWLI can be calculated from information obtained from records and does not require site visits. It can be used to create various scenarios at a train control centre and determine the workload for different sets of variables.

Based on the process followed in this study, the MWLI could be adapted for other control and monitoring environments.

5.3 APPLICATION AND IMPLEMENTATION OF THE MENTAL WORKLOAD INDEX

The phenomenon that resulted in weaker correlations between the MWLI and specific physiological measurements, i.e., habituation to homotypic stressors, is very relevant in the train control environment. Usually, normal operating practices prevail. In other words, following the prescribed procedures is what TCOs do most of the time. It is only under abnormal conditions, such as the failure of radio communication equipment or a dramatic increase in train volumes (as is the case when maize is exported to neighbouring countries), that TCOs are stretched in terms of their capacity to deal with different ways of train control. This is called *abnormal working* in railway terms. Abnormal working forms part of the training of TCOs but is rarely used and only tested when the TCO undergoes refresher training every two years. TCOs therefore become used to, or habituated to, the normal operating procedures. It becomes a habit because it is part of normal daily behaviour.

This study did not focus on abnormal working conditions (or heterotypic stressors), but there is ample evidence that under abnormal conditions (or when *fall back* or abnormal working procedures have to be utilised) the workload becomes dangerously high and errors may result. The work group, all of whom are experienced TCOs, reported that, from personal experience, abnormal working conditions are considered to be very high risk. Everyone in the group reported having had an experience where abnormal working procedures had on occasion led to a near-miss situation in which an accident was averted at the last minute, or an accident was caused due to an error committed by the TCO. This is, however, seldom uncovered when a formal incident inquiry is made.

The reason for the high-risk nature of abnormal working conditions is that these pose a heterotypic stress situation, which adds to the mental workload already experienced by the TCO. Homotypic stressors do not refer to low stress levels but rather to the fact that individuals become used to the type of stressor and therefore do not show physiological responses to the stressor because they do not label the situation as stressful. Under abnormal working it is the novelty of the situation that makes it stressful and therefore compounds the existing mental workload experienced. This increased experience of workload can lead to errors of judgement or incorrect actions.

In recent years several accidents were attributed to abnormal working conditions in the train control environment. None of these occurred in the RTO environment, but all occurred in the colour light signalling environment due to copper cable theft which resulted in the failure of signals. TCOs then have to revert to authorising trains through radio communication (as is done in RTO) but it is not the normal way of working within the colour light environment. Two of the accidents resulted in loss of life and due to the sensitive nature of these accidents and ongoing claims against Spoornet, details may not be provided here.

It is therefore imperative that, in light of this study, cognisance is taken of those train control centres with a high workload in terms of the MWLI. Should there be abnormal conditions, especially in these centres, it should automatically follow that special measures are taken to avert errors that may occur as a result of overload.

5.4 APPLICATION OF A CRITICAL MENTAL WORKLOAD LEVEL

The unavailability of documented research that deals with the practical application of mental workload measurements in terms of a specific cut-off number for safe operations is problematic because very few benchmarks exist. A concept that could be an appropriate application in this context is the *workload redline* concept referred to by De Waard (1996).

In recent years, the question, “How much workload is too much?”, has received increased attention. In an applied setting such as traffic research, the workload redline could prove to be a very useful concept as the consequences of too much workload within the driving environment can be very serious. A proposed model of mental workload, task performance and demands provides the rationale for understanding the workload redline (De Waard, 1996).

A relationship between task demand and task performance has been described by Meister (1976). Meister defined three regions, regions A, B and C. Region A is characterised by low operator workload with high performance. An increase in demands does not lead to performance decrements. In region B the level of performance declines with increased task demands. So, region B is the region where performance decreases with increases in demand and workload. In region C extreme levels of load have diminished performance to a minimum level, and performance remains at this minimum level with further increases in demand (see figure 5.1).

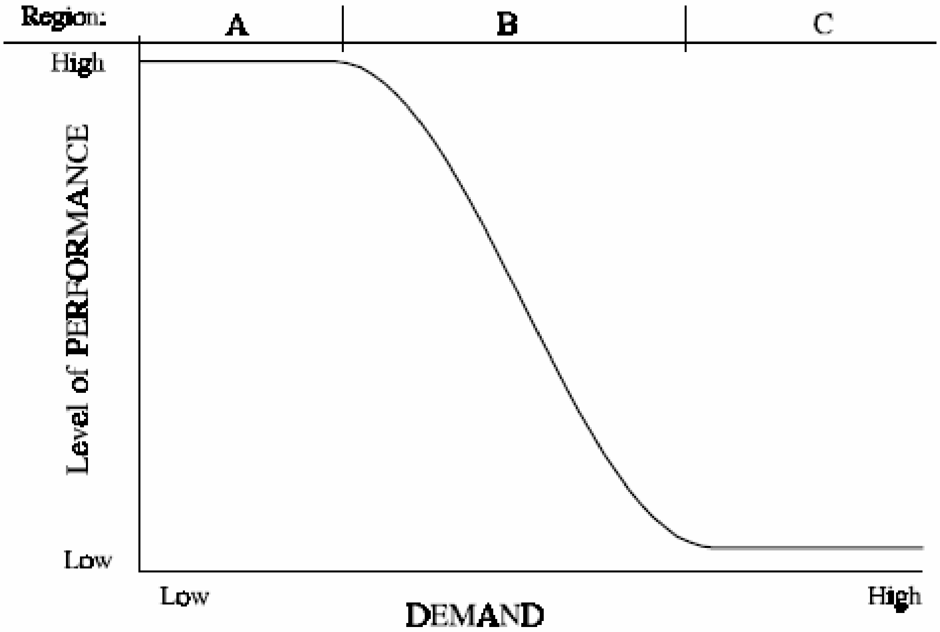


Figure 5.1: Hypothetical relationship between demand and performance (based on Meister,1976)

According to this model, also referred to as the region model, a primary-task workload measure, i.e. a measure of performance, will only be sensitive to variations in levels of workload in region B. In region A performance remains stable and is independent of variations in demand, while in region C performance remains at a minimum level, independent of demand. Other measures, such as self-report (or subjective) measures of workload, may be sensitive in region B and may clearly reveal overload in the C-region, while they need not be sensitive in region A. While extreme levels of load that result in overload can be situated in the C region, it is not clear where the domain of underload is situated (De Waard, 1996).

This model could be completed with the addition of a deactivation or a D region at the far left end. The effects of monotonous tasks, for example, are situated in the D region. These are low-demand tasks that can result in increases in task difficulty and workload as a result of a reduction in capacity. In case of boredom, for example, a reduction in capacity requires that a larger proportion of the capacity be used for performance of the same task, thus increasing mental workload (Meijman and O'Hanlon, 1984). By means of the addition of the D region the complete inverted-U is split into four regions, the D, A, B and C regions (see Figure 5.2).

The issue of too much workload is usually referred to as the determination of a *workload redline* (Reid and Colle, 1988, Wierwille and Eggemeier, 1993). When trying to tackle the determination of a redline there is a need to first decide upon the context of what constitutes ‘too much’. Degraded performance may indicate too much workload, but the effects of personal wellbeing may prove equally valid. Preliminary work on workload redline puts this line at the transition from region A to B (see Figure 5.2).

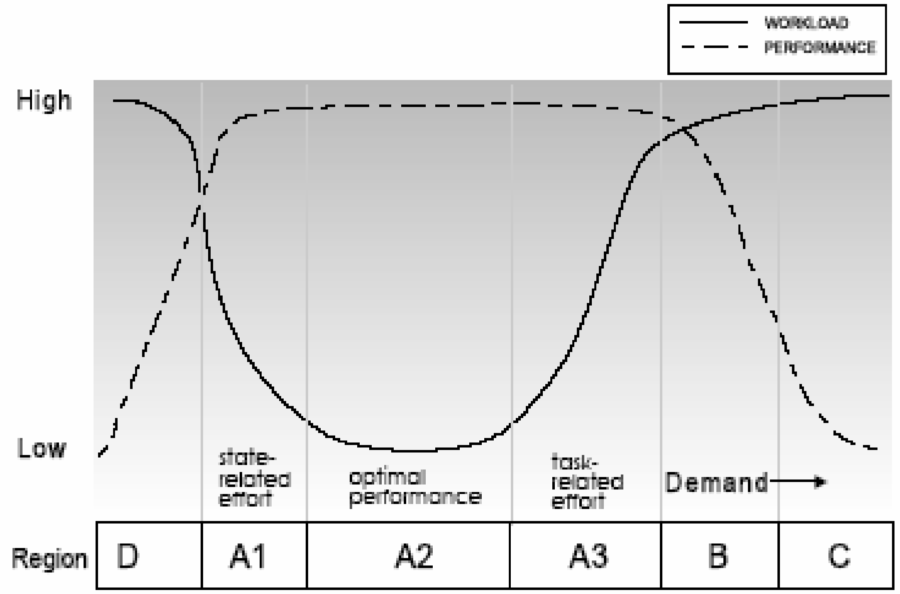


Figure 5.2: Workload and performance in 6 regions (De Waard, 1996)

Reid and Colle (1988) related just-detectable performance decrements to self-report ratings, and this workload rating was used to designate the absolute workload redline. The point of a just-detectable performance decrement occurs at the transition from region A to B. While it is clear that performance measures themselves have defined the A region, it may be useful to split the A-region into three parts. In the middle part, region A2, operators can easily cope with task demands and performance remains at a stable level with increases in demand without increased effort. In the A3 region, however, performance measures still do not show a decline, but operators are only able to maintain the level of performance with increased effort. Temporary compensation by the exertion of effort in region A3 is one of the advantages of human flexibility and is not critical. If, however, continuous effort is required to maintain performance, or if peak loads occur frequently, this can lead to stress, and an unhealthy situation that has to be avoided. This is particularly true if operators have no control over the situation. It may therefore be more useful to put a workload redline at the transition from region A2 to A3 instead of at the transition from region A (A3) to B. In this way, the term *workload redline* remains related to *workload* instead of to primary-task performance breakdown.

A similar situation exists at the region that is to the right of the D-region, region A1. Here, for instance, monotony starts to affect the operator's state, but by 'trying harder', i.e. by the investment of effort, the primary-task performance level is not yet affected. A second workload redline then arises at the transition from region A2 to A1, where operators are effectively counteracting a reduced operator state. When effort investment is no longer effective, the D-region is entered where performance is affected. When demand increases, starting from the optimal operator state in region A2, the operator's capability of (effort) compensation will be exceeded at a certain moment and a transition from the A3 to the B region takes place. In the B region performance is affected and at the moment that it deteriorates to a minimum level the C region is entered (De Waard, 1996).

In region D (*D* for *deactivation*) the operator's state is affected. In region A2 performance is optimal. The operator can easily cope with the task requirements and reach an adequate, self-set level of performance. In the regions A1 and A3 performance remains unaffected but the operator has to exert effort to preserve an undisturbed performance level. In region B this is no longer possible and performance declines, while in region C performance is at a minimum level: the operator is overloaded. (De Waard, 1996)

If the workload redline is not determined by the point at which performance measures start to deteriorate, but is determined by the point at which region A2 is departed, then performance (primary task) measures alone are by definition not sufficient to determine whether the load is unacceptable. Nevertheless, performance measures remain indispensable in redline research to determine whether workload is in the A region. Again, this is an argument in favour of the use of multiple measurements.

One of the aspects of workload measures that is emphasised in workload redline is the use of absolute versus relative measures. Traditionally, relative measures have been used. With relative measures, task performance, self-reports and physiological measures during baseline performance are compared with the same measures during performance of the task or system that is being evaluated. Some authors claim that absolute measures are required for workload redline (e.g., Wierwille and Eggemeier, 1993). So far, critical values on the SWAT rating-scale have been proposed only by Reid and Colle (1988). However, their mention of the critical SWAT value of 40 refers to the point at which performance begins to be affected (the transition from region A to B). Such a workload redline is a *primary-task workload margin*. This margin is defined as a critical level at which the (primary) task has to be performed. Beyond that point, primary-task performance is affected. Although performance margins can be successfully determined, an absolute criterion for workload itself, i.e. the critical value of a measure denoting that region A2 has been left is, according to De Waard (1996), not feasible. The reason for this is that workload is a relative measure; it is the proportion of the capacity that is allocated for task performance. The amount of resources allocated does not depend only upon task demands, but depends also on capability or willingness of the operators to handle the demands.

The conceptual problems of a workload redline become very prominent in applied settings. In traffic, for instance, the capabilities of the population who drive vehicles vary to a great extent. Novice drivers have to allocate more resources for task performance than experienced drivers. Similar differences in capability exist between young and elderly drivers. Consequently, for the same tasks individuals have their own workload redline. In spite of the problems associated with redline definition, an approach that includes primary-task performance margins as they relate to the cost of maintaining performance is useful in any applied field of workload assessment. Self-report scales and performance measures (for the A to B region shift) are probably the most promising measure groups for this. Physiological indices that are opposed to baseline measurements can be very useful in assessing operator effort and the cost of performance maintenance (De Waard, 1996).

It is proposed that, typically, operating in the C region would be dangerous and that a redline standard should be set at a lower level: at the transition from Region A to Region

B. In most circumstances aviation systems should be operated in Region A (Colle and Reid, 2005).

De Waard (1996) the correctness of putting the redline at the point at which performance is affected and suggested as an alternative the point at which effortful, compensatory processes are initiated. For this, the combination of performance measures with physiology and/or self-report measures could provide a clearer picture of mental workload. Critical levels of measures of mental workload are, however, not attainable because mental workload itself is a relative measure. The resources that operators are willing or capable of allocating to task performance differ between individuals. This makes a *redline* that can be defined according to a critical level in a measure of mental workload an impossibility.

The *workload redline* is the only attempt at determining a quantifiable cut-off, or critical, level for mental workload that could be found in the literature. From the above discussion it can be concluded that this is indeed not a simple, if at all possible, task. This approach however provides a useful framework for conducting further research of this concept, which would add value to applied mental workload settings such as traffic control (air and train), driving tasks and other monitoring-task settings.

Practical application of MWLI in terms of a specific cut-off number for safe operations needs further investigation.

CHAPTER 6

CONCLUSIONS

In closing, the aims of the study as well as the results obtained are integrated and a balanced view of the utility of the index is provided. The limitations of the study are discussed and recommendations for further research are offered.

6.1 SUMMARY AND CONCLUDING REMARKS

The primary aim of this research project was to develop an objective measure of mental workload for classifying train control centres. In order to fulfil this aim, the definition and concept *mental workload* was studied. The various definitions provided by researchers were put forward and the different components of workload were discussed and contextualised by the model developed by Meshkati (1988).

The study of the existing body of knowledge often proved contradictory, if not confusing. The clear and consistent message, however, is that there is no empirical technique which can be used to accept or reject a single, simplified definition of workload. Workload is therefore acknowledged as a multidimensional construct and it is recognised that different definitions contribute to the perspective of workload.

As far as mental workload measurement techniques are concerned, the categories of workload measurement were identified and the various techniques representative of the categories were discussed. The review of the research indicated that the workload imposed on operators led to changes that manifested in different ways. These were grouped as changes in performance, changes in subjective experience and physiological changes.

The biggest limitation in the reported literature has been identified as the fact that there are no specific guidelines for the application of workload measurement techniques. In addition to this, no specific guidelines exist with respect to the practical implementation of the techniques and the potential value for the train control environment and rail safety.

The Mental Workload Index (MWLI) did show promising results in discriminating between the *low* and *high* workload categories of train control centres and, if used correctly, could be applied for such purposes. It does not appear to have discriminating abilities in the medium range of mental workload. Validation studies need to be done to determine the point or number on the index that will activate warnings that the workload in a particular centre is dangerously high. This methodology for measuring workload provides a more valid technique than the previous index that Spoornet has been applying. Even without cut-off levels for safe operations, the MWLI could be used to identify high workload centres that require special attention, such as closer supervision, the deployment of more TCOs and the deployment of suitably qualified TCOs. These actions would be appropriate to mitigate the potential or existing hazards in these centres.

Translating the results of the research into norms for safe operations is not a simple exercise, simply because of the variable nature of human operators within the system. There is no clear-cut evidence in the research results that points to a specific number or figure for a physiological parameter that would represent a dangerously high mental workload level. It can be conclusively deduced that most of the TCOs who participated in this study carry a high allostatic load, which could predict health problems in the long term.

Finally, it can be concluded that there is not one single method or procedure that is ideal for the measurement of mental workload. There are, however, sufficient methodologies available that can be applied, and if properly validated, will give a scientifically defendable and practical index of mental workload.

The MWLI developed in this study is, as far as could be ascertained from the existing body of knowledge, a unique contribution to the field of human mental workload measurement, since such an application model, which utilises only objectively gathered and observed information, does not exist. Interest in this model, while it was still in the process of being developed, has already been shown by the nuclear industry in South Africa.

6.2 LIMITATIONS OF THE STUDY

A limitation of this study, which negatively affected the statistical significance of the results, was the small sample population available to the researcher. The reason for this was that the specific train control system (radio train order), which was the focus of the study, was only used in 36 train control centres. The data was therefore limited to a sample of these centres.

A further limitation was the high cost associated with this kind of research, due to the specialised nature of equipment and the knowledge required to analyse and interpret the data. The cost of research is proportional to the sample size. The project had to be executed within severe cost constraints and the sample for the validation study had to be limited in terms of numbers as well as in terms of geographical spread, in an attempt to limit travelling expenses. The verification field study was therefore limited to a small sample of 20 train control centres. Because the data was collected from a small, non-randomly selected sample of TCOs, it may not be representative of all Spoornet TCOs.

The small sample size has resulted in another limitation. Due to the heterogeneity of the sample it was not possible to randomly select both a control and an experimental group. It was also necessary for the given group to be studied within specified geographical boundaries.

6.3 RECOMMENDATIONS

It is suggested that data should be collected more frequently in order to capture the variation and short-term fluctuations in workload that are inherent to a TCO’s job. Data collected every two hours may be insensitive to these variations. Increasing the frequency of data collection would probably not be acceptable to either management or TCOs. These individuals are conducting safety-critical work, so there is also an increased risk of distracting them from their work if more is asked of them. A medium- to high-fidelity dispatching simulator would be more suitable to this type of research. A simulator would enable researchers to control workload and other conditions of the work environment in order to see their effects on the TCO, but as discussed, several obstacles were encountered when attempting to create a high-fidelity simulator specifically for RTO. This possibility is worth exploring if the research could be undertaken in a collaborative effort with an academic or research institution.

It is further suggested that a validation study of the MWLI be undertaken on a larger sample in order to allow for more statistically significant analyses. Once this has been done, the development of a *workload redline* for TCOs should be considered.

As it would entail experimental conditions, the workload-redline development should be performed under controlled conditions in a simulated environment. An *in situ* approach would be invasive and could potentially pose safety risks for the TCOs. It is suggested that the workload redline is approached from two perspectives, namely, from a performance as well as a health perspective. The following approach is suggested:

- The MWLI factors should be 'loaded' in order to recreate conditions that would result in a high calculated MWLI. Under these conditions the physiological parameters used in the field study should be measured to determine the physiological load placed on the TCO.
- Simultaneously, a measure of performance should be performed. Under simulated circumstances it is possible to develop performance indicators for the primary tasks of TCOs. Alternatively, a secondary-task measure of spare mental capacity could be considered.
- The same group of TCOs could then be subjected to conditions resulting in a low calculated MWLI and the concurrent physiological parameters and performance measure could be measured.
- Incremental increases of the MWLI with its associated physiological and performance measure for the same group could be performed.
- With the assistance of physiology experts a 'healthy' threshold of the physiological parameters could be determined and correlated with the MWLI value.
- Similarly, a corresponding MWLI value for deterioration in performance could be established.
- These MWLI values could then be correlated and a redline MWLI value determined.

Future research could also focus on the validation of the MWLI in other safety-sensitive and control or monitoring environments.

The ultimate value of the MWLI would be determined if a longitudinal study of train control incidents (in the RTO environment) could prove a downward trend.

REFERENCES

Averty, P., Collet, C., Dittmar, A., Athènes, S., and Vernet-Maury, E. (2004). Mental Workload in Air Traffic Control: An Index Constructed from Field Tests, *Aviation, Space and Environmental Medicine* **75**(4): 333–341.

Baker, G.J., Suchday, S. and Kranz, D.S. (2000). Heart disease/attack. In Fink, G. (eds.). *Encyclopaedia of Stress*, iii, Academic Press, San Diego, pp. 326–333.

Breierova, L. and Choudhari, M. (1996). Sensitivity Analysis, *Technical report*, MIT System Dynamics in Education Project, Massachussets Institute of Technology.

Casali, J. and Wierwille, W. (1984). On the measurement of pilot perceptual workload: a comparison of assessment techniques addressing sensitivity and intrusion issues, *Ergonomics* **27**: 1033–1050.

Cilliers, M. (1992). *An empirical investigation into the measurement of Fixed Wing fighter pilot workload*, Doctoral thesis, University of Stellenbosch.

Claassen, N., Hazelhurst, L.T., Koorts, A., Van Tonder, J.A., Pretorius, A., Lemmer, H. and Viljoen, M. (2005). Cortisol, haemodynamic responses and heart rate variability in train control officers over an eight-hour day shift, *Proceedings at the 6th IOHA International Scientific Conference, 19-23 September, 2005,*

Colle, H.A. and Reid, G.B. (2005). Estimating a Mental Workload Redline in a Simulated Air-to-Ground Combat Mission, *The International Journal of Aviation Psychology* **15**(4): 303–319.

Damos, D.L. (1988). Individual Differences in Subjective Estimates of Workload. In Hancock, P.A. & Meshkati, N. (eds.). *Human Mental Workload*, North-Holland, Amsterdam, pp. 231–237.

Desmond, P.A., and Hancock, P.A. (2001). Active and Passive Fatigue States. In P.A. Hancock, P.A. and. Desmond, P.A. (eds.). *Stress, Workload and Fatigue,* Lawrence Erlbaum Associates, London, pp. 455–465.

Devoe, D.B. (1974). An analysis of the job of railroad train dispatcher, *Technical Report No. FRA-ORD-74-37*, U.S. Department of Transportation.

De Waard, D. (1996). *The measurement of drivers' mental workload*, PhD Thesis, Haren, Traffic Research Centre, University of Groningen.

Eggemeier, F.T. (1988). Properties of workload assessment techniques. In Hancock, P.A and Meshkati, N. (eds.). *Human Mental Workload,* North-Holland, Amsterdam, pp. 41–62.

Everly, G.S. and Rosenfeld, R. (1981). *The Nature and Treatment of the Stress Response: A Practical Guide for Clinicians*, Plenum Press, New York.

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

Friedman, M and Rosenman, R. (1974). *Type A behaviour and your heart,* Knopf, New York.

Gaillard, A.W.K. (2001). Stress, Workload, and Fatigue as Three Biobehavioral States: A General Overview. In Hancock, P.A. and Desmond, P.A. (eds.). *Stress, Workload and Fatigue,* Lawrence Erlbaum Associates, London, pp. 623–639.

Gawron, J.G., Schflett, S.G. and Miller, J.C. (1989). Measures of in-flight workload. In Jensen, R.S. (ed.). *Aviation Psychology,* Gower Technical, Brookfield, USA, pp. 240–287.

Gericke, C.A. (1994). *The Modelling of Psychophysiological Stress and the Measurement thereof using the Human Electroencephalogram*, M. Eng. Dissertation, University of Pretoria.

Hancock, P.A. (1988). The Effect of Gender and Time of Day upon the Subjective Estimate of Mental Workload during the Performance of a Simple Task. In Hancock, P.A. & Meshkati, N. (eds.). *Human Mental Workload*, North-Holland, Amsterdam, pp. 239–250.

Hancock, P.A. and Meshkati, N. (eds.) (1988). *Human mental workload, Advances in Psychology*, Elsevier, Amsterdam, p. **52**.

Hancock, P.A., Meshkati, N., Robertson, M.S. and Robertson, M.M. (1985). Physiological reflections of mental workload, *Aviation, Space and Environmental Medicine* **56**: 1110–1114.

Hancock, P.A., Rahimi, M., Mihaly, T. and Meshkati, N. (1988). A Bibliographic Listing of Mental Workload Research. In Hancock, P.A. and Meshkati, N. (eds.), *Human Mental Workload,* Elsevier, Amsterdam, pp. 329–333.

Hankins, T.C., and Wilson, G.F. (1998). A Comparison of heart rate, eye activity, EEG and subjective measures of pilot mental workload during flight, *Aviation, Space and Environmental Medicine* **69**(4): 360–367.

Hart, S.G. (1986). Theory and Measurement of Human Workload, *Human Productivity and Enhancement* **1**: 396–445.

Hart, S.G. and Bortolussi, M.R. (1984). Pilot errors as a source of workload, *Human Factors* **26**: 545–556.

Hart, S.G., and Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In Hancock, P.A. & Meshkati, N. (eds.). *Human Mental Workload,* Elsevier, Amsterdam, pp. 139–184.

Hilburn, B. and Jorna, P.G.A.M. (2001). Workload and Air Traffic Control. In Hancock, P.A. & Desmond, P.A. (eds.), *Stress, Workload and Fatigue,* Lawrence Erlbaum Associates, London, pp. 384–394.

Hyyppä, M.T., Aunola, S., Lahtela, K., Lathi, R. and Marniemi, J. (1983). Psychoneuroendocrine responses to mental load in an achievement-oriented task, *Ergonomics* **26**: 1155–1162.

Jex, H.R., (1988). Measuring Mental Workload: Problems, Progress and Promises. In Hancock, P.A. and Meshkati, N. (eds.). *Human Mental Workload*, Elsevier, Amsterdam, pp. 5–40.

Johanssen, G. (1979). Workload and workload measurement. In Moray, N. (ed.). *Mental Workload: Its theory and measurement,* Plenum Press, New York, pp. 3–12.

Kalsbeek, J.W.H. (1971). Standards of acceptable loads in ATC tasks, *Ergonomics* **14**: 641–650.

Kaplan, J.R. (2000). Primate models, cardiovascular disease. In Fink, G. (ed.). *Encyclopaedia of Stress,* Academic Press, San Diego, **3**, pp. 230–236.

Knipling, R.R. and Wierwille, W.W. (1994). Vehicle-based drowsy driver detection: Current status and future prospects, *Paper Delivered at the IVHS America Fourth Annual Meeting, Atlanta, April, 1994.*

Kramer, A.F. (1991). Mental workload: A review of recent papers. In Damos, D. L. (ed.). *Multiple Task Performance,* Taylor and Francis, London, pp. 279–328.

Kroemer, K.H.E., Kroemer, H.B.and Kroemer-Elbert, K.E. (1994). *Ergonomics: How to design for ease and efficiency*, Prentice-Hall Inc., New Jersey.

Kumashiro, M. (2005). Practical measurement of psychophysiological functions for determining workloads. In Wilson, J.R. and Corlett, N. (eds.). *Evaluation of Human Work,* (3 edn.), Taylor and Francis, London, pp. 605–627.

Lacey, J.I. and Lacey, B.C. (1958). The relationship of resting autonomic activity to motor impulsivity. *The Brain and Human Behavior,* Vol 36, Williams and Wilkins, Baltimore.

Megaw, T. (2005). The definition and measurement of mental workload. In Wilson, J.R. and Corlett, N. (eds.). *Evaluation of Human Work,* (3 edn.), Taylor and Francis, London, pp. 525–551.

Meijman, T.F. and O'Hanlon, J.F. (1984). Workload. An introduction to psychological theories and measurement methods. In Drenth, P.J.D., Thierry, H., Willems, P.J. and de Wolff, C.J. (eds.). *Handbook of Work and Organizational Psychology,* Wiley, New York, pp. 257–288.

Meister, D. (1971). *Human Factors: Theory and Practice,* Wiley Intersciences, New York.

Meister, D. (1976). *Behavioral foundations of system development*, Wiley, New York.

Meister, D. (1985). *Behavioral Analysis and Measurment Methods,* John Wiley and Sons, New York.

Meshkati, N. (1988). Heart rate variability and mental workload assessment. In Hancock, P.A. and Meshkati, N. (eds.), *Human Mental Workload*, North-Holland, Amsterdam, pp.101–116.

Meshkati, N., Hancock, P.A., Rahimi, M. and Dawes, S.M. (1995). Techniques in mental workload assessment. In Wilson, J.R. and Corlett, E.N. (eds.), *Evaluation of Human Work: A Practical Ergonomics Methodology*, (2 edn.), Taylor and Francis, London, pp. 749–782.

Meshkati, N. and Loewenthal, A. (1988). The effects of individual differences in information processing behaviour on experiencing mental workload and perceived task difficulty: A preliminary investigation. In Hancock, P.A. and Meshkati, N. (eds.), *Human Mental Workload*, North-Holland, Amsterdam, pp. 269–288.

Meshkati, N. and Loewenthal, A. (1988). An Eclectic and critical review of four primary mental workload assessment methods: A Guide for developing a comprehensive model. In Hancock, P.A. and Meshkati, N. (eds.), *Human Mental Workload*, Elsevier Science Publishers, North-Holland, Amsterdam, pp. 251–268.

Moray, N. (1979). Mental Workload – Its theory and measurement, *Proceedings of the NATO Symposium on Theory and Measurement of Mental Workload, 30 August – 6 September,1977 sponsored by the NATO Special Program Panel on Human Factors, Mati, Greece.* Plenum Press (Published in coordination with NATO Scientific Affairs Division), New York.

Moray, N. (1982). Subjective Mental Workload, *Human Factors* **24**(1): 25–40.

Moray, N. (1984). Mental Workload. *Proceedings of the 1984 International Conference on Occupational Ergonomics, 1984.*

Mulder, G. (1979). Mental Load, Mental Effort and Attention. In Moray, N. (ed.), *Mental Workload: Its theory and measurement,* Plenum Press, New York, pp. 299–326.

Mulder, L.J.M. and Mulder, G. (1987). Cardiovascular reactivity and mental workload. In Kitney, R.I. & Rompelman, O. (eds.). *The Beat-by-beat Investigation of Cardiovascular Function,* Clarendon Press, Oxford, pp. 216–253.

Nunes, A. (2003). The Impact of Automation Use on the Mental Model: Findings from the Air Traffic Control Domain, *Proceedings of the 47th Meeting of the Human Factors and Ergonomics Society,* Santa Monica, CA.

Nygren, T.W. (1991). Psychometric Properties of Subjective Workload Measurement Techniques, *Human Factors* **33**: 17–33.

O'Donnell, R.D. and Eggemeier, F.T. (1986). Workload assessment methodology. In Boff, K., Kaufman, L. & Thomas, J. (eds.). *Handbook of Perception and Human*

Performance, Cognitive Processes and Performance, Vol. II, Wiley and Sons, Inc, New York, pp. 42.1–42.49.

Parasuraman, R. and Hancock, P.A. (2001). Adaptive Control of Mental Workload. In Hancock, P.A. and Desmond, P.A. (eds.). *Stress, Workload and Fatigue,* Lawrence Erlbaum Associates, London, pp. 305–320.

Pew, R.W. (1979). Secondary tasks and workload measurement. In Moray, N. (ed.). *Mental Workload: Its Theory and Measurement,* Plenum Press, New York, pp. 23–28.

Pickup, L., Wilson, J.R., Sharples, S., Norris, B., Clarke, T. and Young, M.S. (2005). Fundamental examination of mental workload in the rail industry, *Theoretical Issues in Ergonomics Science* **6**(6): 463–482.

Popkin, S., Gertler, J. and Reinach, S. (2001). A Preliminary Examination of Railroad Dispatcher Workload, Stress and Fatigue, *Technical Report No. DOT/FRA/ORD-01-08*, Department of Transportation,Washington, DC, U.S.

Pulat, B.M. (1992). *Fundamentals of Industrial Ergonomics,* Prentice-Hall Inc., New Jersey, USA.

Rahimi, M. and Wierwille, W.W. (1982). Evaluation of the sensitivity and intrusion of workload estimation techniques in piloting tasks emphasizing mediational activity, *Proceedings of the IEEE International Conference on Cybernetics and Society*, 593–597.

Rail Safety and Standards Board (2005). Human Factors. [*Research Catalogue CD Rom.*]

Rasmussen, J. (1979). Reflections on the concept of Operator Workload. In Moray, N. (ed.). *Mental Workload: Its theory and measurement,* Plenum Press, New York, pp. 29–40.

Reid, G.B. and Colle, H.A. (1988). Critical SWAT values for predicting operator overload. *Proceedings of the Human Factors Society 32nd annual meeting, Human Factors and Ergonomics Society, Santa Monica, CA,* pp. 1414–1418.

Reid, G.B. and Nygren, T.E. (1988). The subjective workload assessment technique: a scaling procedure for measuring mental workload. In Hancock, P.A. and Meshkati, N. (eds.). *Human Mental Workload*, North-Holland, Amsterdam, pp. 185–218.

Reinach, S. (2001). Preliminary Development of a Railroad Dispatcher Taskload Assessment Tool: Identification of Tasks and Data Collection Methods. *Technical Report,* U.S. Department of Transportation.

Reinach, S. (2006). Toward the development of a performance model of railroad dispatching. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting,* pp. 2042–2046.

Rubio, S., Diaz, E., Martin, J, and Puente, J.M. (2004). Evaluation of Subjective Mental Workload: A comparison of SWAT, NASA-TLX, and Workload Profile Methods, *Applied Psychology: An International Review*, **53** (1), 61–86.

Saltelli, A., Chan, K., Scott, E.M. (2000). *Sensitivity Analysis.* [S.I.:s.n.].

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

Senders, J. (1979). Axiomatic model of workload. In Moray, N. (ed.). *Mental workload: Its theory and measurement,* Plenum Press, New York, pp. 263–267.

Steptoe, A. (1981). *Psychological Factors in Cardiovascular Disorders*, Academic Press Inc., London.

Tsang, P.S. and Johnson, W.W. (1987). Automation: Changes in cognitive demands and mental workload. In Jensen, R.S. (ed.). *Proceedings of the Fourth* International

Symposium on Aviation, The Ohio State University and the Association of Aviation Psychologists Columbus, OH, pp. 616–622.

Ursin, H. and Ursin, R. (1979). Physiological indicators of mental workload. In Moray, N. (ed.). *Mental Workload: Its Theory and Measurement*, Plenum Press, New York, pp. 349–366.

Van Tonder, J.A. (1999). *Certificate in Industrial Ergonomics,* University of Pretoria, Pretoria. [Course Manual.]

Veltman, J.A. and Gaillard, A.W.K. (1998). Physiological workload reactions to increasing levels of task difficulty, *Ergonomics* **41**: 656–669.

Vidulich, M.A. (1988). The Cognitive Psychology of Subjective Mental Workload. In Hancock, P.A. & Meshkati, N. (eds.). *Human Mental Workload*, North-Holland, Amsterdam, pp. 219–229.

Wickens, C.D. (1992). *Engineering Psychology and Human Performance* (2 edn.), Harper-Collins, New York.

Wickens, C.D. (2001). Workload and Situation Awareness. In Hancock, P.A. and Desmond, P.A. (eds.). *Stress, Workload and Fatigue*. Lawrence Erlbaum Associates, London, pp. 443–450.

Wickens, C.D. and Hollands, J.G. (2000). *Engineering Psychology and Human Performance,* (3 edn.), Prentice-Hall Inc., New York.

Wickens, C.D., and Yeh, Y.Y. (1983). The Disassociation of Subjective Ratings and Performance: A Multiple resources Approach. *Proceedings of the Human Factors 27th Annual Meeting, Human Resources Society, Santa Monica, CA,* pp. 244–288.

Wierwille, W. W. (1979). Physiological Measures of Aircrew Mental Workload, *Human Factors*, **21**(5): 575–593.

Wierwille, W. W. (1988). Important remaining issues in Mental Workload Estimation. In Hancock, P.A. & Meshkati, N. (eds.). *Human Mental Workload,* North-Holland, Amsterdam, pp. 315–327.

Wierwille, W.W. and Casali, J.G. (1983a). A valid rating scale for global mental workload measurement applications. *Proceedings of the 27th Annual Meeting of the Human Factors Society, Human Factors Society, Santa Monica, CA,* pp. 129–133.

Wierwille, W.W. and Eggemeier, F.T. (1993). Recommendation for mental workload measurement in a test and evaluation environment, *Human Factors* **35**: 263–281.

Wierwille, W.W. and Gutmann, J.C. (1978). Comparison of Primary and Secondary Task Measures as a Function of Simulated Vehicle Dynamics and Driving Conditions. *Human Factors* **20**(2): 233–244.

Wierwille, W.W., Rahimi, M. and Casali, J.G. (1985). Evaluation of sixteen measures of mental workload using a simulated flight task emphasizing mediational activity, *Human Factors* **27**: 499–502.

Williges, R.C. and Wierwille, W. W. (1979). Behavioral measures of air crew mental workload, *Human Factors* **21**: 549–574.

Wilson, J.R. and Corlett, E.N. (1995). *Evaluation of Human Work: A Practical Ergonomics Methodology,* (2 edn.), Taylor and Francis, London.

Wilson, J.R. and Corlett, N. (2005). *Evaluation of Human Work,* (3 edn.), Taylor and Francis, London.

Wilson, G.F. and O'Donnell, R.D. (1988). Measurement of operator workload with the Neuropsychological Workload Test Battery. In Hancock, P.A. and Meshkati, N. (eds.). *Human Mental Workload,* North-Holland, Amsterdam, pp. 63–100.

Woo, M.A., Stevenson, W.G., Moser, D.K., Trelease, R.B. and Harper, R.M. (1992). Patterns of beat-to-beat heart rate variability in advanced heart failure, *American Heart Journal* **1992** (March): 704–710.

Xie, B. and Salvendy, G. (2000). Review and reappraisal of modelling and predicting mental workload in single- and multi-task environments, *Work and Stress*, **14**.1: 74–99.

APPENDIX A

APPENDIX A: Cohen's Perceived Stress Questionnaire

ID: …………………… Date ……/………/20…..
Station: ……………… Time of day ……….h………min Shift time: ………………

The following questions ask you about your feelings and thoughts DURING THE LAST MONTH. In each case, you will be asked to indicate how often you felt or thought a certain way. Although some of the questions are similar, there are differences between then and you should treat each one as a separate question.

		How Often In the last Month? (Circle your answer)				
		Never	Almost Never	Some-times	Fairly Often	Very Often
1	In the last month, how often have you been upset because of something that happened unexpectedly?	0	1	2	3	4
2	In the last month, how often have you felt that you were unable to control the important things in your life?	0	1	2	3	4
3	In the last month, how often have you felt nervous and "stress"?	0	1	2	3	4
4	In the last month, how often have you dealt successfully with irritating life hassles?	0	1	2	3	4
5	In the last month, how often have you felt that you were effectively coping with important changes that were occurring in your life?	0	1	2	3	4
6	In the last month, how often have you felt confident about your ability to handle your personal problems?	0	1	2	3	4
7	In the last month, how often have you felt that things were going your way?	0	1	2	3	4
8	In the last month, how often have you found that you could not cope with all the things that you had to do?	0	1	2	3	4
9	In the last month, how often have you been able to control irritation in your life?	0	1	2	3	4
10	In the last month, how often have you felt that you were on top of things?	0	1	2	3	4
11	In the last month, how often have you been angered because of things that happened that were outside of your control?	0	1	2	3	4
12	In the last month, how often have you found yourself thinking about things that you have to accomplish?	0	1	2	3	4
13	In the last month, how often have you been able to control the way you spend your time?	0	1	2	3	4
14	In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?	0	1	2	3	4

The following questions ask you about your feelings and thoughts DURING THE LAST SHIFT. In each case, you will be asked to indicate how often you felt or thought a certain way. Although some of the questions are similar, there are differences between then and you should treat each one as a separate question.

1	How stressful did you experience the shift? 1 = No stress 5 = Very stressful	1	2	3	4	5
2	Was this shift more or less stressful than usual?	More stressful Less stressful				
3	Was the shift unpleasant /terrible/ very bad?	Yes No				
4	Was the shift pleasant/nice?	Yes No				

Reason(s) why the shift was pleasant/unpleasant:

APPENDIX B

APPENDIX B: TIMELINE ANALYSIS (EXAMPLE TEMPLATE)

SPOORNET PROJECT: TCO MENTAL WORKLOAD

Timeline Analysis

		Frequency per 15 minute intervals															
Activity	**Total Frequency**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Task preparation for shift																	
Establish trains scheduled for shift																	
Plot planned train movements on train plan																	
Radio communications																	
Plan train movements																	
Update train plan with real time information																	
Issue authorisation																	
Telephone conversations																	
Direct enquiries (in office)																	
Data capturing (ETA,ETD)																	
Write report(s)																	
Personal (bathroom, coffee)																	
Other information																	
Day of the week																	
Shift	From:	To:															
Subjective experience of fatigue		Very tired				Tired				A little tired							

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Addendum to Appendices C and D

The two reports attached as Appendix C and Appendix D has confidentiality requirements set by the writers of the reports. These requirements have been waived by the writers of the reports (see the attached letter), on condition that Spoornet grants permission thereto.

Permission was obtained form Spoornet to use all the data related to this study for the purposes of a PhD dissertation, provided that the following condition is met:
“No individuals may be identified”.

Both reports are therefore attached to the dissertation and may be made available in the public domain.

The role of the candidate in the pilot and validation study was the design of the research parameters and approach, setting the requirements of the protocols to be used, as well as the criteria for testing and the desired outcomes.

Adele Pretorius
30 November 2007

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

AD DESTINATUM PERSEQUOR

University of Pretoria

P.O. Box 2034 Pretoria 0001
Tel (012) 319 2140 Fax (012) 321 1679
mviljoen@medic.up.ac.za
Department of Physiology
School of Medicine
Faculty of Health Sciences
University of Pretoria

28 November 2007

To whom it may concern

Inclusion of the report, *Validation of Spoornet Mental Workload Index against an Allostatic Load Index,* as appendix to the PhD thesis of Adele Pretorius

We hereby declare that we, Prof M Viljoen and Dr N Claassen, don’t have any objection should the candidate wish to attach the above report to her thesis, given that Spoornet grants written permission thereto.

M. Viljoen

Prof M Viljoen

inter

Adele Pretorius
Assistant Manager
Human Resources (Staffing)

Reference:

Telephone: 77-35374 Fax: 77-35378 E-mail: Date: 26 April 20

D.Litt.etPhil – Utilisation of Mental Workload Project Data

Approval is granted to utilise the above data for a dissertation required in fulfillment of the above degree.

The following conditions will apply :

- No individuals to be identified;
- No publication of data (beyond dissertation) without further approval; and
- Two copies of dissertation to be supplied to Spoornet library.

Regards,

JAPIE BENADE
ACTING GENERAL MANAGER
(HUMAN RESOURCES)

Human Resources

PO Box 2179, Joubert Park, South Africa, 2044.

Spoornet is a division of Transnet Limited registration No. 90/00900/06

SPOORNET

APPENDIX C

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Pilot study to investigate the stress levels in train control officers

Prof M Viljoen en Dr N Claassen

September 2001

CONFIDENTIAL

Introduction

The stress response can be defined as the physiological and psychological changes, which occur as a result of the impact of stressors. It is important to realize that many factors can influence the effect of the environmental factors such as work conditions, i.e., will determine whether conditions will be perceived as stressors or not. Such factors include interpersonal differences such as previous experiences, conditioning, genetics and age, as well as temporal differences where the individual's response to a specific stressor could vary at different occasions. Two factors of major importance in determining whether an individual will perceive a situation as stressful or not are a) the appetitive-aversive qualities and b) the degree to which the individual feels himself in control of the situation. Both these factors will not only influence the cognitive, but also the physiological response to the situation.

The stress response can be measured in terms of acute stress or as the long-term effects of chronic exposure to stressful situations. The latter is generally referred to as the allostatic load and is assessed by measuring the negative impact on a number of physiological factors that represent a fair reflection of the general physical condition (1). The influence of the allostatic load on the cognitive and emotive functions can also be assessed by means of a battery of psychological tests.

Stressors can generally be divided into physical stressors and psychological stressors - with psychological stressors referring to all situations that can induce cognitive and emotional activation states. The stress response resulting from psychological stressors is, with minor variations, relatively non-specific with regard to the original stimulus and the neuroendocrine response. Activation of the two major stress axes, i.e., the sympatho-adrenomedullary axis (SAM-axis) and the hypothalamo-adrenocortical axis (HPA-axis) is, to varying degrees, considered to be characteristic of all psychologically-induced activation states. The activity of these axes can be determined by their hormonal/neurotransmitter status, by changes in the physiological processes under control of the axes or, in the case of the evaluation of the effect of chronic stress, by the relevant pathophysiological changes in the body. In theory the neurohormonal shifts would be the best indicators of acute stress, but in case of the SAM-axis it is often better to ascertain the sympathoadrenomedullary activational state by measuring physiological activities controlled by the system. The best functional parameters for this purpose are probably the registration of heart rate, in case of acute stress, and blood pressure in case of chronic stress determinations. In the case of the HPA-axis the functional alterations require a significant amount of time for full expression and acute variations in activational state are therefore not well reflected in such determinations. The acute reaction is thus best assessed by measuring the level of the major target hormone, cortisol.

A number of important variations in the response of the two major stress axes to psychological stressors have, however, recently been observed. It has for instance been observed that heart rate can fairly consistently be found to increase with mental effort while blood pressure is often very little influence with mental effort without physical involvement like speaking or moving around. Major differences in the stress response have also been reported between a high activational state coupled to aversive emotional experiences and that with neutral or appetitive emotions. An interesting observation, yet to be further substantiated, is the difference observed in the response in the activational state of working in order to avoid a negative outcome and that of an activational state of working for a monatory award (2)

The parameters to be assessed are shortly reviewed at this stage in an attempt to avoid an unnecessary discourse during the discussion of the results.

Blood pressure as stress indicator

Increased blood pressure is a major complication of stressful life styles. However, measurements of blood pressure as an indicator of acute stress often do not yield the information required to assess the magnitude of the stress response (3). The reason for this is multi-factorial, but the major confounding factors include the fact

a) that intermittent behavioural stress often leads to sustained, potentially pathogenic increases in both systolic and diastolic pressures. The major underlying physiological mechanism is the stress-induced increases in plasma lipids – a phenomenon closely related to the development of atherosclerosis.

b) that relatively chronic psycho-social stress negatively influence endothelium-dependent vasodilatory responses

Heart rate as a stress indicator

Heart rate is largely the product of direct sympathetic nervous system responses, or indirect sympathetic system responses (via activation of the adrenal medulla and the circulating catecholamine pool), and of the involvement of neurohormonal factors of the HPA-axis in the regulation of the sympathetic system. In contrast to blood pressure, the sensitivity of heart rate to stress-induced central nervous system activation is generally considered to render it a justified index of acute psychological stress. The rationale for this lies in the fact that transient changes in heart rate can be detected superimposed on chronic, pathophysiological increased heart rates. Consistent high baseline heart rate values are usually the result of factors like genetics, cardiovascular problems or anaemia and hardly ever the direct effect of psychological stress. Transient, stress-induced increases in pulse rate therefore offer a reliable index of acute conditions of stress. According to the cardiovascular reactivity hypothesis cardiovascular reactivity to chronic stressor exposure contributes to the development of hypertension, myocardial infarctation and stroke. While heart rate fluctuations are probably amongst the best indicators of acute stress, it appears not to be of prognostic value – this in contrast to blood pressure which is a poor reflection of transient stress but a good prognostic parameter of eventual pathophysiology (4).

A number of factors should be considered before heart rate variations are summarily taken as the status quo of the cognitive-emotive status of the individual. The two most important of these include:

a) the manner in which the heart rate assessments are performed. Heart rate values obtained by palpitation of a second person over an artery can have several disadvantages. The first disadvantage entails the fact that the manual collections of heart rate counts can only be done intermittently and a continuous recording would thus not be available. Important shifts in heart rate may therefore be missed. The second problem is that of accuracy of the palpitation counts – anyone ever involved with this procedure over an extended period of time is well aware of the pitfalls. The third, and probably the major, confounding factor, is the influence of the proximity of the observer on the psychology and therefore not only on the heart rate, but also on the ability of the test person to continue with normal activities required for task performance. It is thus of paramount importance that additional confounding aspects, such as manual assessments of the heart rate response, should not be introduced into the already problem-riddle field of research.

b) the type of emotional and cognitive response to a stressor.
The typical flight-or fight response generally leads to in increase in blood pressure – the underlying physiological mechanisms are well known. It was shown that heart rate generally increases with most types of emotional experiences, with the exception of disgust. Heart rate increases are generally to be expected with cues for punishment or reward, especially reward. The response is naturally influence by the perception of coping or the inability to cope. This is especially true if some physical action is involved, During a period of anticipatory attention, before the individual goes into action, it is a fairly common occurrence to find that the heart rate is slowing down. This fact is supported by results obtained in experiments where the stressor or stimulus comprised computer games that contain periods of anticipation. It is a common error in stress research to expect the same type of pulse rate reaction form all types of emotion. Examples that substantiate the variation in heart rate responses to different types of psychological stressors are seen in

- ☐ *The startle response* sometimes referred to as the orientating response and by some seen as the “what is it response?” This alerting-related bradycardia can be seen in conditions when the physical fight-or-flight response is not appropriate – when physical activity does not form part of the defense pattern.

- The cardiovascular changes associated with the *conditioned emotional response* to a stressor are also known not to conform to expectations. It may result in either bradycardia or tacycardia (5).

In the present experimental test results the heart rates were determined by two independent electronically monitored recording apparatus.

Cortisol as indicator of stress

In response to stress the hypothalamus would secrete corticotropin releasing hormone which would stimulate the anterior pituitary to release adrenocorticotropin which in turn would stimulate the adrenal cortex to release the glucocorticoid hormone, cortisol. Together these structures and their hormones represent the HPA-axis. Cortisol is secreted in a diurnal rhythm that is reflected in the plasma concentration of cortisol. Cortisol concentrations in the peripheral blood, and in secretions like saliva, reach a peak in the early morning hours between 06:00 and 08:00, falls progressively towards noon, with a small rise occurring just before lunch, to eventually reach the lowest level around 20:00 to 21:00 hours. This rhythm can be disturbed by a number of factors. Of importance for this study is the fact that it may be disturbed by shift work where the natural day/night time pattern of being awake and being asleep is not observed (6).

Activation of the cardiovascular responses, particularly heart rate can, as discussed in a previous paragraph, occur during mental and emotional stressful tasks, regardless of the negativity or positivity of the emotional tone. In the case of cortisol the emotional quality of the task would appear to determine the magnitude of the response. It is becoming ever more evident that potently aversive situations can lead to prompt and substantial increases in cortisol secretion. It therefore seems feasible to see the cortisol response as a distinguishing feature of distressing events (7).

The brief for this pilot study was: a) to compare the magnitude of acute stress induced by high work loads to that induced by low work load shifts, in terms of physiological variables, and b) to assess whether any correlation could tentatively be observed between the physiological indices of acute stress and that of the work loads. The latter in order to assist in the quest for finding guidelines in the development of a formula, based on the frequency of various activities, by which the various centers can be rated

Methods

Methods

Four male train control officers situated at [redacted] Train Control center were used in this study. After explaining all relevant procedures to the train control officers, they signed a volunteer informed consent form. Their anthropometric data are depicted in table 1. Table 2 depicts the work shifts of the train control officers under investigation at [redacted] train control center during the period of investigation. Subject 1 and 2 worked at the "low" (L) and "high" (H) stress work loads respectively for a time period of one month preceding the evaluation. Their shift schedule followed a seven-day day-shift period (Sunday to Sunday), followed by a seven-day night shift period. S3 and S4 alternated on a day-to-day basis between the L and H work loads. The same weekly shift regime was followed as described for S1 and S2. Due to personnel shortages, S4 worked additional hours from 12:00 – 06:00 the night before the low stress evaluation.

In order to get the maximum experimental information from the four available train control officers, each officer was tested over 2 full shifts. Two were on the same load shift and two on a cross over basis between high and low shifts.

Measurements:

Before the evaluation started the subjects completed the Cohen's perceived stress scale (PSS) to assess their non-specific, appraised stress during the last month.. This questionnaire consists of 14 items of which 7 are positively formulated (eg, "In the last month, how often have you felt that things go your way?") and 7 are negatively formulated (e.g. "In the last month, how often have you felt that you were unable to control the important things in your life?") (Appendix A).

After each day of evaluation the subjects completed a questionnaire to assess their feelings they have experienced during the previous shift (Appendix B).

Heart rate was measured with Polar heart rate monitors at a 15 second interval (Polar Electro). Each subject was issued with a chest harness consisting of two "dry" electrodes, connected to a miniature radio signal transmitter. Electrical activity of the heart is processed and transmitted to a receiver to allow data sampling of the test subject's heart rate during exposure.

Blood pressure, i.e., diastolic (DBP) and systolic (SBP) blood pressure, were measured with a digital electronic blood pressure meter (ALP K2, model DS-125D, Japan). This device was calibrated against a calibrated mercury blood pressure manometer. Two consecutive blood pressure measurements were taken at the beginning of each work session and every two hours there after during the shift. The average of the two measurements is reported.

Saliva was collected in clean collection test tubes, after each blood pressure measurement. Salivary collections were sampled at 06:00, 08:00, 10:00, 12:00, 14:00, 16:00 and 17:30, and stored on ice after collection. The collected saliva was centrifuged at 1000 g for 10 minutes. An aliquot of the supernatant was transferred into two 1.5 ml eppendorf tubes and stored at –20°C until analysis. Free salivary cortisol was determined with a Salivary Cortisol ELISA kit (SLV-2930, DRG Instruments GmbH, Frauenbergstrasse, Marburg, Germany). The principle of the cortisol ELISA kit is based on the competition principle and microplate separation. An unknown amount of salivary cortisol and a fixed amount of cortisol conjugated with horse-radish peroxidase compete for binding sites of a polyclonal cortisol antiserum coated onto the wells. After one hour of incubation, the microtiterplate was washed to stop the competition reaction. The absorbance of each well was determined at 450 nm. The free cortisol levels are inversely proportional to the optical density measured.

A time line analysis, a measurement of mental work load, was recorded during the 12 hour work shift. The same observer was responsible for the time line analysis at specific work loads respectively. The main activities recorded were the updating time schedules, train orders, number of trains, radio communication and telecommunications.

Table 1: Anthropometric data of the train control officers taking part in the study at the [redacted] train control center.

Subject	Age (y)	Mass (kg)	Length (m)	Body mass index (kg.m^{-2})
S1	31	85.5	1.74	28.2
S2	41	124	1.73	41.4
S3	37	64	1.60	25.1
S4	39	95	1.71	32.4
Average	37.0	92.1	1.70	31.8
Standard deviation	4.3	24.9	0.07	7.1

Table 2: Work shifts of train control officers under investigation at [redacted] train control center.

Subject	Date	Time of shift	Work load classification
S1	13 June 2001	06:00 – 18:00	“Low”
S1	15 June 2001	06:00 – 18:00	“Low”
S2	13 June 2001	06:00 – 18:00	“High”
S2	15 June 2001	06:00 – 18:00	“High”
S3	19 June 2001	06:00 – 18:00	“High”
S3	20 June 2001	06:00 – 18:00	“Low”
S4	19 June 2001	06:00 – 18:00	“Low”
S4	20 June 2001	06:00 – 18:00	“High”

Results

Tables 3 to 13 summarize the data measured and calculated from four train control officers. The tables include individual and average work load of each work load level.

Table 3: Perceived stress evaluation score (PSS), number of trains and train orders issued and post shift evaluation of train control officers at the [redacted] train control centre. (L = “Low stress, H = “High stress, R = Repeat, ND = Not determined)

					Post shift evaluation remarks questionnaire			
Person ID	Workstation ID	PSS	Trains	Train orders	How stressful was the shift 1 = No stress 5 = Very stressful	Was the shift more or less stressful than usual	Was the shift unpleasant	Was the shift pleasant
1	L	11	12	14	3	Less stressful	No	Yes
1	LR		8	8	3	Less stressful	No	Yes
Mean	L	11	10	11				
Stdev	L		2.8	4.2				
2	H	20	12	22	1	Less stressful	No	Yes
2	HR		16	24	1	Less stressful	No	Yes
Average	H	20	14	23				
Stdev	H		2.8	1.4				
3	L	28	5	5	ND	ND	ND	ND
3	H		13	34	3	More stressful	No	Yes
4	L	19	5	5	1	Less stressful	No	Yes
4	H		20	22	ND	ND	ND	ND
Average	L	23.5	5.0	5.0				
Stdev	L		0.0	0.0				
Average	H	23.5	16.5	28.0				
Stdev	H		4.9	8.5				

Table 4: Systolic blood pressure (mm.Hg^{-1}) of train control officers at the [redacted] train control centre during a twelve hour shift. (L = "Low stress, H = "High stress, R = Repeat)

		Time (hh:mm)						
Person ID	Workstation ID	06:00	08:00	10:00	12:00	14:00	16:00	17:30
1	L	138	127	131	125.5	132	116.5	128
1	LR	130	122	126	127	115	139	135
Average	L	134	124.5	128.5	126.25	123.5	127.75	131.5
Stdev	L	5.7	3.5	3.5	1.1	12.0	15.9	4.9
2	H	140	140	149	153	147	147	154.5
2	HR	134.5	143	135	154	136.5	138	146.5
Average	H	137.25	141.5	142	153.5	141.75	142.5	150.5
Stdev	H	3.9	2.1	9.9	0.7	7.4	6.4	5.7
3	L	141.5	143.5	135	141.5	145.5	153.5	137
3	H	158	152	147	139	157	147	154
4	L	124	120.5	123	133	130.5	129	128.5
4	H	133.5	129	135.5	118	122	128	130
Average	L	132.8	132.0	129.0	137.3	138.0	141.3	132.8
Stdev	L	12.4	16.3	8.5	6.0	10.6	17.3	6.0
Average	H	145.8	140.5	141.3	128.5	139.5	137.5	142.0
Stdev	H	17.3	16.3	8.1	14.8	24.7	13.4	17.0

Table 5: Diastolic blood pressure (mm.Hg^{-1}) of train control officers at the [redacted] train control centre during a twelve hour shift. (L = “Low stress, H = “High stress, R = Repeat)

		Time (hh:mm)						
Person ID	Workstation ID	06:00	08:00	10:00	12:00	14:00	16:00	17:30
1	L	93	90	88	86	87	82	93.5
1	LR	98	91	93.5	95	89	104.5	85
Average	L	95.5	90.5	90.75	90.5	88	93.25	89.25
Stdev	L	3.5	0.7	3.9	6.4	1.4	15.9	6.0
2	H	86	115	88.5	109.5	99.5	94	90.5
2	HR	100	94.5	96	97	91	86.5	113.5
Average	H	93	104.75	92.25	103.25	95.25	90.25	102
Stdev	H	9.9	14.5	5.3	8.8	6.0	5.3	16.3
3	L	97.5	94.5	92.5	102	86.5	102	97.5
3	H	107.5	127.5	91.5	96	102	101.5	112.5
4	L	87	95	95.5	85	83	86.5	87.5
4	H	97	89.5	90	88.5	83.5	91	104
Average	L	92.3	94.8	94.0	93.5	84.8	94.3	92.5
Stdev	L	7.4	0.4	2.1	12.0	2.5	11.0	7.1
Average	H	102.3	108.5	90.8	92.3	92.8	96.3	108.3
Stdev	H	7.4	26.9	1.1	5.3	13.1	7.4	6.0

Table 6: Pulse pressure (mm.Hg^{-1}) of train control officers at the [redacted] train control centre during a twelve hour shift. (L = "Low stress, H = "High stress, R = Repeat)

		Time (hh:mm)						
Person ID	Workstation ID	06:00	08:00	10:00	12:00	14:00	16:00	17:30
1	L	45.0	37.0	43.0	39.5	45.0	34.5	34.5
1	LR	32.0	31.0	32.5	32.0	26.0	34.5	50.0
Average	L	38.5	34.0	37.8	35.8	35.5	34.5	42.25
Stdev	L	9.2	4.2	7.4	5.3	13.4	0.0	11.0
2	H	54.0	25.0	60.5	43.5	47.5	53.0	64.0
2	HR	34.5	48.5	39.0	57.0	45.5	51.5	33.0
Average	H	44.3	36.8	49.8	50.3	46.5	52.3	48.5
Stdev	H	13.8	16.6	15.2	9.5	1.4	1.1	21.9
3	L	44.0	49.0	42.5	39.5	59.0	51.5	39.5
3	H	50.5	24.5	55.5	43	55	45.5	41.5
4	L	37.0	25.5	27.5	48.0	47.5	42.5	41.0
4	H	36.5	39.5	45.5	29.5	38.5	37.0	26.0
Average	L	40.5	37.3	35.0	43.8	53.3	47.0	40.3
Stdev	L	4.9	16.6	10.6	6.0	8.1	6.4	1.1
Average	H	43.5	32.0	50.5	36.3	46.8	41.3	33.8
Stdev	H	9.9	10.6	7.1	9.5	11.7	6.0	11.0

Table 7: Mean blood pressure (mm.Hg^{-1}) of train control officers at the [redacted] train control centre during a twelve hour shift. (L = “Low stress, H = “High stress, R = Repeat)

		Time (hh:mm)						
Person ID	Workstation ID	06:00	08:00	10:00	12:00	14:00	16:00	17:30
1	L	107.9	102.2	102.2	99.0	101.9	93.4	104.9
1	LR	108.6	101.2	104.2	105.6	97.6	115.9	101.5
Average	L	108.2	101.7	103.2	102.3	99.7	104.6	103.2
Stdev	L	0.5	0.7	1.4	4.6	3.0	15.9	2.4
2	H	103.8	123.3	108.5	123.9	115.2	111.5	111.6
2	HR	111.4	110.5	108.9	115.8	106.0	103.5	124.4
Average	H	107.6	116.9	108.7	119.8	110.6	107.5	118.0
Stdev	H	5.3	9.0	0.3	5.7	6.5	5.7	9.0
3	L	112.0	110.7	106.5	115.0	106.0	119.0	110.5
3	H	124.2	135.6	109.8	110.2	120.2	116.5	126.2
4	L	99.2	103.4	104.6	100.8	98.7	100.5	101.0
4	H	109.0	102.5	105.0	98.2	96.2	103.2	112.6
Average	L	105.6	107.0	105.6	107.9	102.3	109.8	105.8
Stdev	L	9.1	5.1	1.4	10.0	5.2	13.1	6.7
Average	H	116.6	119.1	107.4	104.2	108.2	109.9	119.4
Stdev	H	10.7	23.4	3.4	8.5	16.9	9.4	9.6

Table 8: Heart rate (beats.min^{-1}) of train control officers at the [redacted] train control centre during a twelve hour shift. (L = “Low stress, H = “High stress, R = Repeat)

		Time (hh:mm)						
Person ID	Workstation ID	06:00	08:00	10:00	12:00	14:00	16:00	17:30
1	L	65.0	47.0	45.0	51.5	50.0	55.0	49.0
1	LR	68.5	52.5	50.5	54	65	68.5	55.5
Average	L	66.8	49.8	47.8	52.8	57.5	61.8	52.3
Stdev	L	2.5	3.9	3.9	1.7	10.6	9.5	4.6
2	H	87.0	79.0	89.0	96.0	84.0	82.0	92.0
2	HR	80.0	83.0	81.5	96.5	80.0	86.0	84.5
Average	H	83.5	81.0	85.3	96.3	82.0	84.0	88.3
Stdev	H	4.9	2.8	5.3	0.4	2.8	2.8	5.3
3	L	72.0	73.0	58.0	72.5	66.5	79.0	61.0
3	H	75.5	71.5	79.0	78.0	71.0	64.0	72.5
4	L	91.5	81.0	82.5	82.0	88.5	90.5	87.5
4	H	92.5	84.0	95.5	80.0	77.5	75.5	87.5
Average	L	81.8	77.0	70.3	77.3	77.5	84.8	74.3
Stdev	L	13.8	5.7	17.3	6.7	15.6	8.1	18.7
Average	H	84.0	77.8	87.3	79.0	74.3	69.8	80.0
Stdev	H	12.0	8.8	11.7	1.4	4.6	8.1	10.6

Table 9: Cortisol (ng.mL^{-1}) of train control officers at the [redacted] train control centre during a twelve hour shift. (L = “Low stress, H = “High stress, R = Repeat)

		Time (hh:mm)						
Person ID	Workstation ID	06:00	08:00	10:00	12:00	14:00	16:00	17:30
1	L	10.5	5.5	4.5	4.75	4.5	4.0	2.5
1	LR	11.0	4.0	5.5	5.5	3.5	7.5	5.0
Average	L	10.8	4.8	5.0	5.1	4.0	5.8	3.8
Stdev	L	0.4	1.1	0.7	0.5	0.7	2.5	1.8
2	H	32.0	11.5	7.5	6.5	7.0	7.5	9.0
2	HR	12.0	10.0	8.0	6.0	5.0	4.0	5.0
Average	H	22.0	10.8	7.8	6.3	6.0	5.8	7.0
Stdev	H	14.1	1.1	0.4	0.4	1.4	2.5	2.8
3	L	14.5	8.5	12.5	11.0	12.0	10.0	6.5
3	H	12.5	12.0	11.5	10.5	7.5	2.5	6.5
4	L	13.0	16.0	10.5	12.0	14.0	7.5	3.0
4	H	16.0	14.5	7.0	7.0	10.0	3.5	2.5
Average	L	13.8	12.3	11.5	11.5	13.0	8.8	4.8
Stdev	L	1.1	5.3	1.4	0.7	1.4	1.8	2.5
Average	H	14.3	13.3	9.3	8.8	8.8	3.0	4.5
Stdev	H	2.5	1.8	3.2	2.5	1.8	0.7	2.8

Table 10: Radio communication (Number per 2 hours) of train control officers at the [redacted] train control centre during a twelve hour shift. (L = "Low stress, H = "High stress, R = Repeat)

		Time (hh:mm)					
Person ID	Workstation ID	06:00 – 08:00	08:00 – 10:00	10:00 – 12:00	12:00 – 14:00	14:00 – 16:00	16:00 – 17:30
1	L	17	13	8	16	7	4
1	LR	12	4	16	0	22	5
Average	L	14.5	8.5	12.0	8.0	14.5	4.5
Stdev	L	3.5	6.4	5.7	11.3	10.6	0.7
2	H	30	33	46	36	51	15
2	HR	51	67	58	62	64	45
Average	H	40.5	50.0	52.0	49.0	57.5	30.0
Stdev	H	14.8	24.0	8.5	18.4	9.2	21.2
3	L	0	0	19	8	17	17
3	H	49	59	66	56	84	34
4	L	0	2	23	14	10	12
4	H	47	43	56	60	57	32
Average	L	0.0	1.0	21.0	11.0	13.5	14.5
Stdev	L	0.0	1.4	2.8	4.2	4.9	3.5
Average	H	48.0	51.0	61.0	58.0	70.5	33.0
Stdev	H	1.4	11.3	7.1	2.8	19.1	1.4

Table 11: Telecommunications (Number per 2 hours) of train control officers at the [redacted] train control centre during a twelve hour shift. (L = "Low stress, H = "High stress, R = Repeat)

		Time (hh:mm)					
Person ID	Workstation ID	06:00 – 08:00	08:00 – 10:00	10:00 – 12:00	12:00 – 14:00	14:00 – 16:00	16:00 – 17:30
1	L	1	2	5	2	3	0
1	LR	0	5	7	0	6	0
Average	L	0.5	3.5	6.0	1.0	4.5	0.0
Stdev	L	0.7	2.1	1.4	1.4	2.1	0
2	H	6	15	9	6	5	1
2	HR	5	7	15	6	4	3
Average	H	5.5	11.0	12.0	6.0	4.5	2.0
Stdev	H	0.7	5.7	4.2	0	0.7	1.4
3	L	2	3	9	1	4	2
3	H	11	8	8	3	4	5
4	L	3	0	5	6	2	9
4	H	6	2	7	8	6	1
Average	L	2.5	1.5	7.0	3.5	3.0	5.5
Stdev	L	0.7	2.1	2.8	3.5	1.4	4.9
Average	H	8.5	5.0	7.5	5.5	5.0	3.0
Stdev	H	3.5	4.2	0.7	3.5	1.4	2.8

Table 12: Routine notations (Number per 2 hours) of train control officers at the [redacted] train control centre during a twelve hour shift. (L = “Low stress, H = “High stress, R = Repeat)

		Time (hh:mm)					
Person ID	Workstation ID	06:00 – 08:00	08:00 – 10:00	10:00 – 12:00	12:00 – 14:00	14:00 – 16:00	16:00 – 17:30
1	L	0	3	3	6	5	3
1	LR	2	0	11	0	11	5
Average	L	1.0	1.5	7.0	3.0	8.0	4.0
Stdev	L	1.4	2.1	5.7	4.2	4.2	1.4
2	H	15	30	39	25	35	9
2	HR	31	40	37	33	42	30
Average	H	23.0	35.0	38.0	29.0	38.5	19.5
Stdev	H	11.3	7.1	1.4	5.7	4.9	14.8
3	L	0	0	15	5	10	12
3	H	24	33	39	40	66	23
4	L	0	0	14	6	1	11
4	H	37	30	36	39	44	19
Average	L	0.0	0.0	14.5	5.5	5.5	11.5
Stdev	L	0.0	0.0	0.7	0.7	6.4	0.7
Average	H	30.5	31.5	37.5	39.5	55.0	21.0
Stdev	H	9.2	2.1	2.1	0.7	15.6	2.8

Table 13: Correlation between selected physiological variables with radio communication, telephonic inquiries and routine notations over the 12 hour exposure time.

Physiological variable	Radio communication		Telephonic communication		Routine notations	
	r	p	r	p	r	p
Low stress						
Heart rate	0.0373	0.6649	0.1576	0.0659	0.0823	0.3392
Cortisol	-0.1250	0.6212	0.0764	0.7633	-0.0421	0.8684
SBP	0.2355	0.3469	0.1221	0.6294	00.3859	0.1138
DBP	0.0122	0.9617	0.0762	0.7638	0.1876	0.4559
Mean arterial BP	0.1334	0.5978	0.1154	0.6485	0.3294	0.1819
Pulse Pressure	0.2419	0.3336	0.0800	0.7522	0.2875	0.2474
High stress						
Heart rate (2 hour value)	-0.5164	0.0282	-0.0081	0.9745	-0.4343	0.0717
Heart rate (Polar)	0.2290	0.3607	-0.0534	0.8334	0.2552	0.3068
Cortisol	-0.0648	0.7984	0.4221	0.0810	-0.2375	0.3427
SBP	-0.0764	0.7631	0.0749	0.7677	-0.0883	0.7276
DBP	-0.2921	0.2395	0.0976	0.6999	-0.3580	0.1447
Mean arterial BP	-0.2366	0.3446	0.0979	0.6991	-0.2878	0.2467
Pulse Pressure	0.2362	0.3454	-0.0180	0.9436	0.2960	0.2330
Low & High stress						
Heart rate (2 hour value)	0.3312	0.0485	0.3396	0.0427	0.3314	0.0483
Heart rate (Polar)	0.4550	0.0053	0.2880	0.0885	0.4610	0.0047
Cortisol	-0.1299	0.4501	0.1752	0.3068	-0.1668	0.3308
SBP	0.3482	0.0374	0.2396	0.1593	0.3463	0.0385
DBP	0.2367	0.1645	0.2298	0.1775	0.1990	0.2447
Mean arterial BP	0.3071	0.0685	0.2572	0.1300	0.2804	0.0976
Pulse Pressure	0.1951	0.2542	0.0664	0.7002	0.2310	0.1753

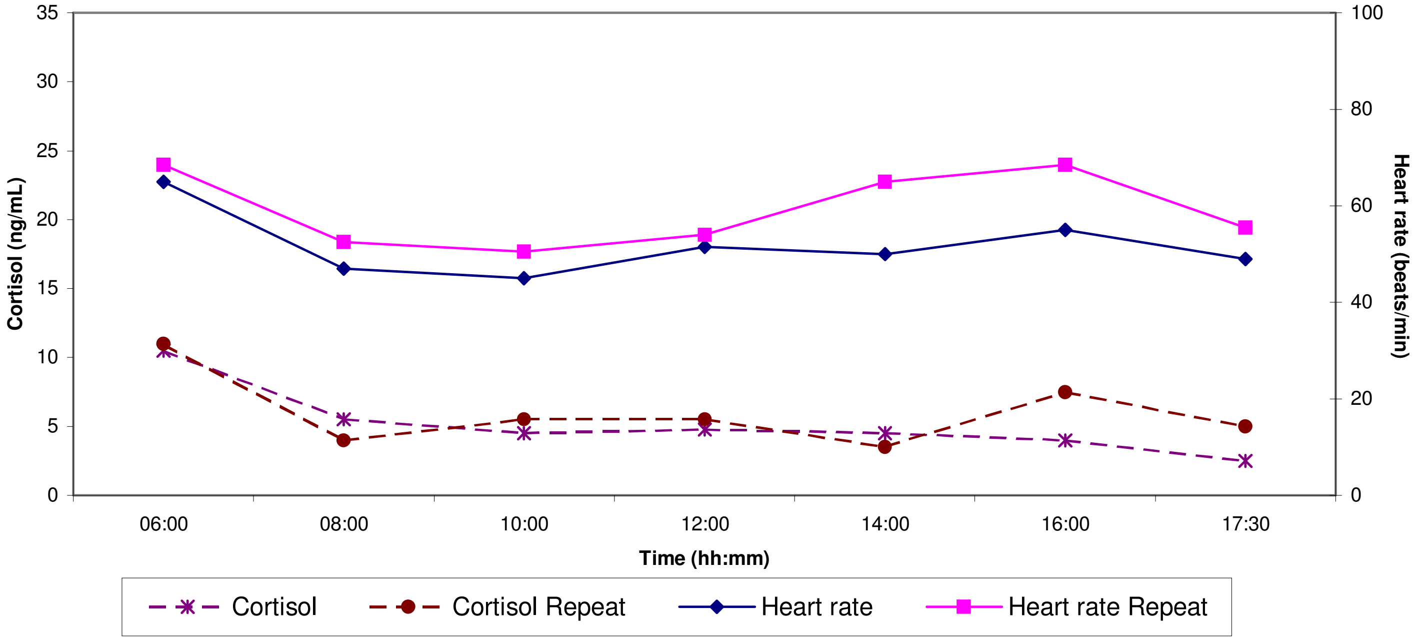


Figure 1: Cortisol and heart rate responses of S1 to a low work load

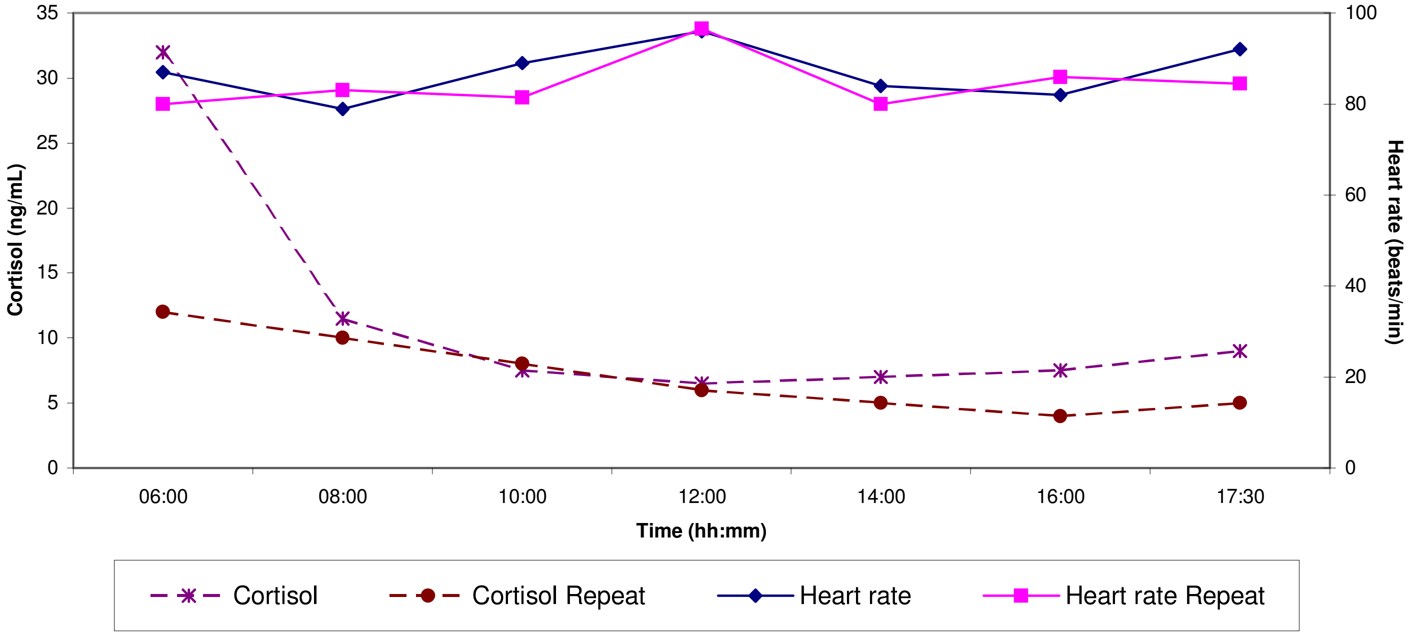


Figure 2: Cortisol and heart rate responses of S2 to a high work load

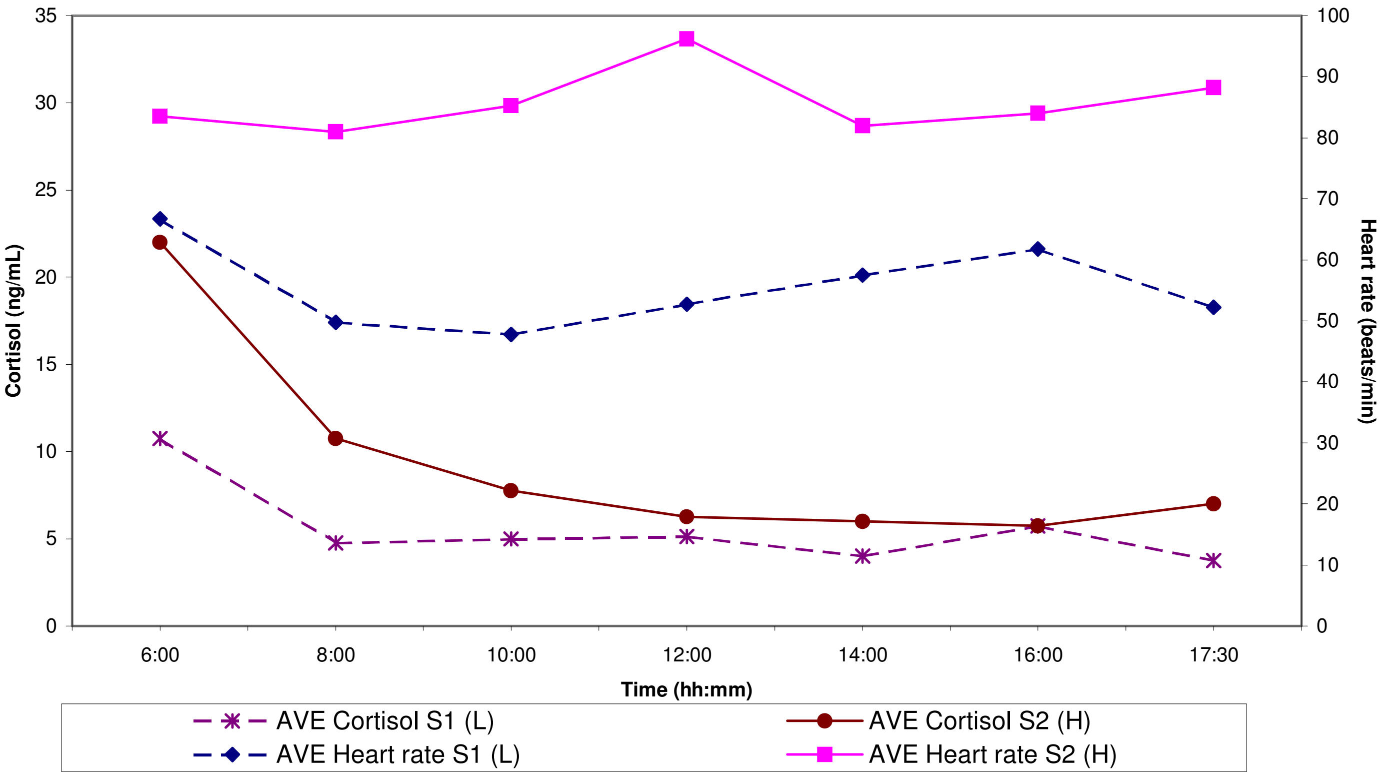


Figure 3 Mean cortisol and heart rate responses of S1 and S2 at a low and high work loads respectively.

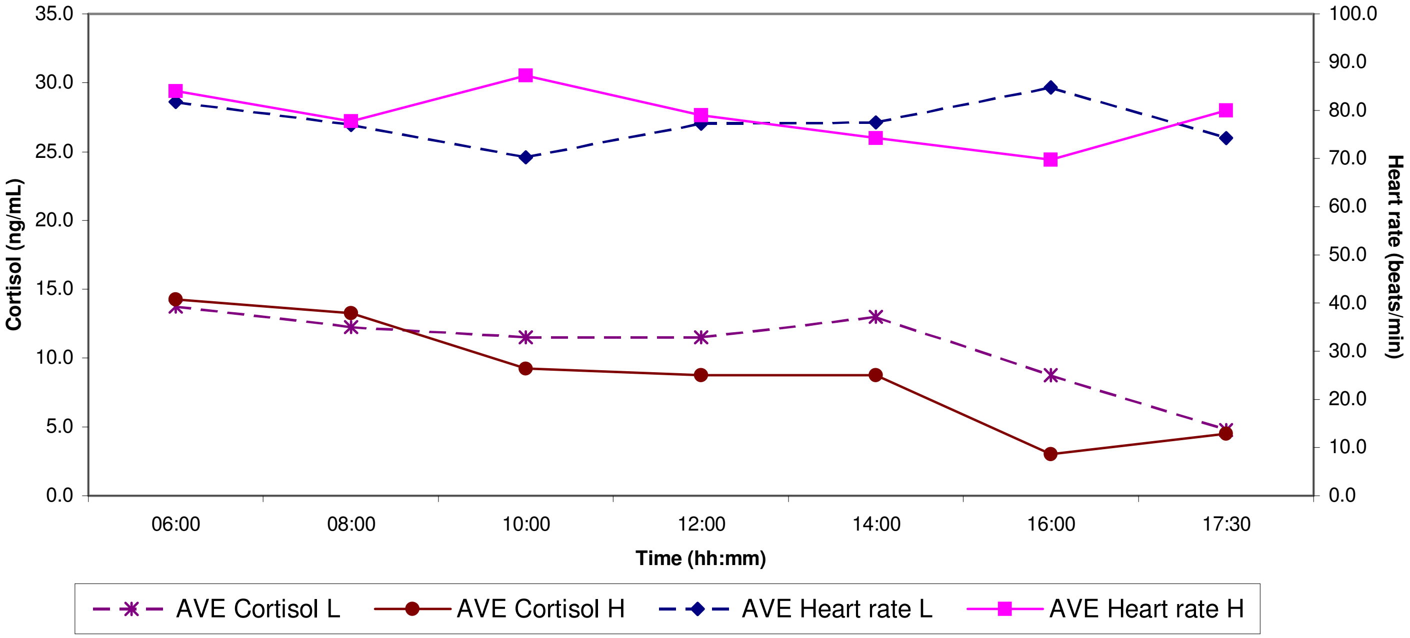


Figure 4: Mean cortisol and heart rate responses of S3 and S4 at a low and high work loads respectively.

Discussion

The brief for this pilot study was to a) compare the magnitude of acute stress induced by low work load shifts to that induced by high work load shifts, in terms of physiological variables and b) to assess whether any correlation could tentatively be observed between the physiological indices of acute stress and that of the work loads.

Theoretically, the effect work stress in situations like this, should be assessed by determining the allostatic load in a statistically valid number of individuals who were consistently on high work load shifts for an extended period of time and comparing their values to those of the same number of individuals consistently on low load shifts for an extended period of time. Factors like hours per shift, number of shifts per month and several other variables should then also be taken in consideration. In this study the number of workers, the work stations and the work conditions were beyond the control of the investigators. For this reason the current study should be seen as a pilot study. Nevertheless, a number of indications and important guidelines were observed.

The results are presented as Part A, which deals with the stress comparisons at different work loads, and Part B, which deals with the correlations between the numerical assessment of work load and the magnitude of the physiological stress response.

To facilitate the reading of this report the reader is at this stage presented with the outlay of the discussion:

Part A: A comparison between the physiological stress responses at high work loads to that at low work loads

a) Difference in work intensity between high and low load shifts
b) Reproducibility of results
c) Differences in the responses to high work load and to low work load when all the high load values were pooled into one group and all the low load values into another.

d) Differences in the responses to high work load and to low work load when the values of individuals were analysed.
e) Differences in the responses to high work load and low work load as seen during a cross-over comparison.
f) Conclusions on the rating of work loads in terms of the physiological response

Part B of study: Activity levels or work load versus physiological indicators of stress

a) Activity levels or work load versus physiological indicators of stress for the total group
b) Activity levels or work load versus physiological indicators of stress, observed during low load shifts
c) Activity levels or work load versus physiological indicators of stress observed during high load shifts
d) Conclusions on the correlation between activity scoring and physiological stress response scoring

Part C: Tentative indications of high allostatic loads

Part D: Final conclusions and recommendations

Part A: A comparison between the physiological stress responses at high work loads to that at low work loads

The aim was to examine the stress levels in terms of the physiological response at low work loads compared to that at high work loads. In looking at the data it is imperative to remember that every individual already carried the effects of the allostatic load, a factor which may have a confounding effect on the values.

The discussion to follow is based on the information in the Tables and Figures under the results section.

a) Difference in work intensity between high and low load shifts

To ascertain whether the work loads were indeed different between the presumed high and low load shifts, the number of trains, train orders and the frequency of the three major activities were compared (one-way ANOVA).

:

Number of trains per high load shift: mean = 15.7; SD = 3.8
Number of trains per low load shift: mean: = 6.7; SD = 2.9
Statistical significance of difference: p = 0.0307

Train orders per high load shift: mean = 26.3; SD = 6.7
Train orders per low load shift: mean = 7.0: SD = 3.5
Statistical significance of difference: p = 0.0111

Routine notations of scheduled work per high load shift: mean = 34.1; SD = 11.1
Routine notations of scheduled work per low load shift: mean: = 5.5; SD = 5.2
Statistical significance of difference: p = 0.00001

Telecommunications per high load shift: mean = 6.1; SD = 3.2
Telecommunications per low load shift: mean = 3.4; SD = 2.8
Statistical significance of difference: p = 0.0104

Radiocommunications per high load shift: mean = 51.2; SD = 13.0
Radiocommunications per low load shift: mean = 10.2; SD = 6.9
Statistical significance of difference: p = 0.00001

c) Reproducibility of results

The reproducibility of the determinations over one entire shift was tested by evaluating the values from duplicate shifts by the same individual on which all parameters were measured at 2hrs interval. This was done on one volunteer at a low load station and one at a high load station. The reproducibility of the two main stress measurements, i.e. cortisol levels and heart rate for each can be seen in Figures1-2 of the results section. The first determination of each shift was ignored as it reflected not only the peak of the circadian rhythm but also the physical and psychological activities before the initiation of the work session. The reproducibility between the duplicate shifts was good and no significant difference was found.

d) Differences in the responses to high work load and to low work load when all the high load values were pooled into one group and all the low load values into another.

The mean values over time and the statistical differences between the values for high and for low work loads, obtained when the one-way ANOVA was applied were as follows:

> The mean arterial blood pressure (mmHg) for the high load shifts (112 ± 9.9; n = 18) was significantly higher than that for the low load shifts (105 ± 5.5; n = 18): p = 0.0139 (Kruskal-Wallis)
>
> The diastolic blood pressure (mmHg) for the high load shifts (98 ± 10.5; n = 18) were significantly higher than that for the low load shifts (91 ± 5.4; n = 18): p = 0.0376 (Kruskal-Wallis)
>
> The systolic blood pressures (mmHg) for the high load shifts (140 ± 11.5; n = 18) were significantly higher than for the low load shifts (132 ± 8.8;n = 18): p = 0.0215
>
> No significant difference were found between the high and low load shifts for the salivary cortisol (ng/ml): p = 0.5432

The average heart rate over time (beats/min) for the high load shifts (90 ± 3.1; n = 18) were significantly higher than for the low load shifts (81 ± 10.7; n = 18):
p = 0.0169 (Krushkal-Wallis)

From the above values one would at first glance conclude that the physiological stress response is indeed significantly higher during the high work load shifts. This would be based on the significant differences in the cardiovascular responses – in this case representative of the activation of the SAM-axis. Cortisol levels, a HPA-axis stress indicator is usually expected to be elevated by aversive psychological experiences. According to the results obtained when the values of all determinations were pooled into two activity groups, no difference existed between the appetitive-aversive perceptions of the two work load individuals.

The next step was to analyse the values of the individuals in both groups to test whether the statistical significance of the differences observed for the total group were indeed valid.

e) Differences in the responses to high work load and to low work load when the values of individuals were analysed.

The differences between the response to high and low loads were first statistically analysed on the values obtained from one individual (#2) twice on high, and one individual (#1) twice on low work loads.

In Figure 3 the graphs for the mean cortisol and mean heart rate of the duplicate determinations are presented for both the low and the high work load candidates. The visible differences were confirmed by the results of the statistical analysis (one-way ANOVA) of the data. In comparing the mean over time values obtained from the high load to the values of the low work load shifts it was shown that cortisol values for the 2 high load shifts (mean = 7.25; SD = 1.9; n = 6), when compared to salivary cortisol values for the low load shifts (mean = 4.73; SD = 1.4; n = 6) were significantly higher: p = 0.0117

and

heart rate/min values for the 2 high load shifts (mean = 91; SD = 1.7; n = 6), when compared to average heart rate values for the low load shifts (mean = 68; SD = 3.4; n = 6) were significantly higher: p = 0.00001

This would once again seemed to support the existence of a significantly higher physiologically stress response during high activity shifts. It should, however be kept in mind that the individuals who were each tested twice over their respective work loads were not age, BMI, race and allostatic load matched and the results should therefore be viewed with caution. The non-matching of candidates were the result of the limitations on the availability of matching subjects at the particular venues and the restrictions in the number of candidates imposed on the study. For this kind of comparison to be accepted without reservations one of two improvement to the study design should be made: a) either the experimental procedure should be performed on at least 200 individuals - each, preferably, confined to one type of work load or even better, b) a cross-over study should be performed where the values of each candidate involved in the study are obtained at both high and low work schedules.

The next step in this pilot study was indeed a cross-over assessment where each of two candidates was evaluated on both high and low load shifts.

f) Differences in the responses to high work load and to low work load as seen during a cross-over comparison.

In this comparison the values of two candidate (#3 & #4) were in the first place analysed by comparing the values of the same candidate at high and at low work loads, and in the second place by pooling (calculating the mean) the values over time obtained at the two high load sessions and comparing that to the mean of the two low load sessions. Only the statistics of the two major acute stress indicators are presented here. The other values and raw data can be found in the results section.

Mean over time of the cortisol levels for subject #3 at high work load: 8.4ng/ml (SD = 3.6; n = 6)

Mean over time of the cortisol levels for subject #3 at low work load: 10.1ng/ml (SD = 2.3; n = 6)

No significant difference existed between the series of values: p = 0.3635

Mean over time of the cortisol levels for subject #4 at high work load: 7.4ng/ml (SD = 4.4; n = 6)

Mean over time of the cortisol levels for subject #4 at low work load: 10.5ng/ml (SD = 4.7; n = 6)

No significant difference existed between the series of values: p = 0.2674

The standard deviations for the cortisol values pointed towards large fluctuations in cortisol levels which, when translated into terms of stress, may indicate periods of aversive experiences. It is of interest that the tendency is only found during high load shifts.

Mean over time for the heart rate values for subject #3 at high work load: 88/min (SD = 3.5; n = 6)

Mean over time of the heart rate values for subject #3 at low work load: 84/min (SD = 2.8; n = 6)

The difference bordered on statistical significance and would most probably have been significant with larger experimental groups: p = 0.0656

Mean over time for the heart rate values for subject #4 at high work load: 91/min (SD = 2.9; n = 6)
Mean over time of the heart rate values for subject #4 at low work load: 92/min (SD = 2.6; n = 6)
No statistical difference was found between the two series of values: p = 0.4639 –
in fact, the standard deviations were very small reflecting only minor deviation in the heart rate over both the high and low load shifts

g) Conclusions on the rating of work loads in terms of the physiological response

The results presented in this section of the study confirmed the statistically significant differences between the workloads of the high and low shifts. In evaluating the reproducibility of results it was shown that the physiological responses, obtained during duplicate evaluations at both high and low loads, did not differ significantly and that it should generally not be considered a problem area of the research. In comparing the physiological parameters at high workloads to that at low workloads it was shown that the values were indeed uniformly higher during the high load shifts. However, this may be a reflection of allostatic load rather than acute stress.

Values of individuals were subsequently analysed. The differences were, however, confirmed by comparing the values of one individual tested twice on high load, and one individual tested twice on low load.

To try and eliminate non-matching in personal allostatic loads and other factors, a crossover study was performed with two workers – each doing a low load and a high load. From the results it can with a fair amount of confidence be said that the heart rate response can be an indicator of differences in workload, but that larger experimental groups should be tested to absolutely confirm this. An interesting observation was that high workload does not appear to be experienced as an aversive factor.

Part B of study: Activity levels or work load versus physiological indicators of stress

This part of the study was performed in an attempt to find guidelines for the development of a formula, based on the intensity of the work schedule, which could differentiate between high and low stress station. The information on which this discussion is based can be seen in under results in Table 13.

a) Activity levels or work load versus physiological indicators of stress for the total group

In the present study positive correlations were found between the activity scores and some of the physiological parameters of stress if the values of all experimental subjects were taken into account. The activity levels were assessed by subdividing activity into three subclasses, i.e., radio communications, telecommunications and routine notations of scheduled train traffic and related events. It should be stressed that telecommunications are often the means of communications if there is a breakdown in the standard lines of communications or if special request or enquiries beyond the duties of the worker take place.

The radio-communication scores correlated positively with the mean arterial blood pressure values, the systolic blood pressure and the heart rate

Radio communication versus SBP: $r = 0.3482$; $p = 0.0374$
Radio communication versus Mean arterial BP: $r = 0.3071$; $p = 0.0685$
Radio communication versus HR: $r = 0.3312$; $p = 0.0485$

The validity of the stat recorded heart rate values were tested by means of continuous recordings of heart rates by a polar watch. The validity of the positive correlations just shown was supported by the values of the correlations obtained when the mean values of the continuous recordings were used in the statistical analysis

Radio communication versus HR polar: $r = 0.4550$; $p = 0.0053$

This statistically significant correlation between magnitude of activity and physiological activation was also observed when telecommunications for the total group were compared to heart rate. Again increases in telecommunication correlated positively with increases in heart rate

Telecommunication versus HR: $r = 0.3396$; $p = 0.0427$

Again the validity of the significance was confirmed by alternative heart rate recordings. The correlations between the telecommunications and heart rates as obtained by means of the polar heart rate monitor recordings were

Telecommunication versus HR polar: $r = 0.2880$; $p = 0.0885$

The correlation between activity scoring and physiological activation for the total group was once more confirmed when comparing the activities involved in the routine work such as notation of the schedules.

Correlations as obtained from stat heart rate values automatically recorded by blood pressure monitor were

Routine notations versus SBP: $r = 0.3463$; $p = 0.0385$
Routine notations versus HR: $r = 0.3314$: $p = 0.0483$

When once again testing the validity by polar watch recordings the correlation was also confirmed.

Routine notations versus HR polar: r = 0.4610: p = 0.0047

No statistical significant correlations were found between cortisol and activity scores when the experimental group as a whole were considered.

In analysing the data for the experimental subjects as a total group it can be concluded that a correlation does indeed exist between the activity scoring and the physiological scoring of stress as depicted by cardiovascular responses. The fact that this was found mainly for heart rate and for cardiovascular parameters that involved cardiac function (systolic blood pressure and mean arterial pressure) rather than purely vascular reactivity, and that no correlations could be found between activity and cortisol levels, is of significance. It is known, as discussed in the introduction, that the emotional characteristics of a stress response are of paramount importance for determining the autonomic and endocrine responses. Emotionally positive, activating challenges will result in heightened cardiovascular responses – and in the case of acute stress situations, more specifically acute transient heart rate responses that coincide with the time of stressor application. As long as no significantly aversive reaction is experience during the work related activity-induced stress response, cortisol levels will normally not increase to any significant level. The total picture seen in the response of the group to increases in work activity can thus be summarised by saying firstly that, as would be expected, a good correlation does indeed exist between the work load as measured by activity scores and the degree of psychologically-induced physiological activation. This is a normal phenomenon and would generally occur, even if the activity happened to be recreational or pleasurable in nature. Secondly, and of great importance, is the fact that the transient increases in physiological activation were not accompanied by negative emotional experiences – in other words the workers would not seem to have experience any marked degree of distress as a result of an increase in work load. The latter deduction, based on the results of the objective results, is further confirmed by the subjective score sheets filled in by the workers at the end of each shift.

The next step was to subdivide the experimental values into values obtained during high workloads and that obtained during low work loads.

b) Activity levels or work load versus physiological indicators of stress, observed during low load shifts

No relevant statistical significant correlations were found either for the SAM-axis stress indicators (cardiovascular reactivity) or HPA-axis stress indicators (cortisol) when the values obtained at all low load shifts were pooled. As before the validity of the heart rate recordings were confirmed by a second series of determinations by polar watch.

When the values obtained from individual workers were examined it was noticeable that for one specific individual (#1), the one with generally the lowest basal values in terms of stress indicators, the comparisons between activities involving verbal communications and cortisol levels were just marginally non-significant

Radio communications versus cortisol: $r = 0.7999$; $p = 0.0561$
Telecommunications versus cortisol: $r = 0.7945$; $p = 0.0590$

With a larger n-value it is highly likely that the p-value would have been significant. The conclusion to be reach from this observation is that verbal communication presents a degree of thread to this individual.

c) Activity levels or work load versus physiological indicators of stress observed during high load shifts

In comparing the physiological stress parameter values obtained during the high load shifts, with the exception of one parameter, no correlations could be found between physiological stress indicators and the frequency of the various activities. A statistical significant negative correlation was observed between radio communication and heart rate with the single automatic recording system.

Radio communications versus HR: r = - 0.5164; p = 0.0282

Polar recordings of heart rates were then compared to the frequency of radio communications. These recordings recorded no correlation with radio communication.

Radio communications versus HRpolar: r = 0.2290; p = 0.3607

It is a fairly common occurrence to find that the heart rate is slowing down during a period of anticipatory attention, i.e., before the individual goes into action. The possibility therefore existed that such a phenomenon could have been the cause of the initial negative correlation seen between heart rate and the frequency of radio communications. However, in testing the correlations by using the mean of the continuous recordings the correlation falls away and it become obvious that a high variability in heart rate may be the more feasible explanation. This observation stresses the suspicion that continuous recordings of heart rate give a more accurate reflection and should be the technique of choice.

Indications of a positive correlation were seen between cortisol and the number of telecommunications per high workload shift when all the values obtained over high load shifts were considered. The correlation was only marginally not significant and indications are that with larger experimental groups it would be shown that the handling of telecommunications, superimposed on an already high workload schedule, is the one activity found to have the most aversive quality.

Telecommunications versus cortisol levels: r = 0.4221; p = 0.0810

When examining the values of individual workers the mean values for cortisol of subjects #2 and #3, taken intermittently at predetermined times over their shifts, showed a statistically significant positive correlation with telecommunications. These two workers alternated between high and low work stations and a crossover comparison could therefore be made between their values at high and that at low work loads. It is important to note that no correlation could be found between telecommunication and cortisol levels at low work shifts

High work load(S3&S4): telecommunications versus cortisol: r = 0.6487; r = 0.0225
Low work load (S3&S4): telecommunications versus cortisol: r = -0.1852; r = 0.5645

In these individuals it can, with a fair degree of certainty be said that telecommunications are experienced as an aversive intrusion when already taxed with a high workload.

d) Conclusions on the correlation between activity scoring and physiological stress response scoring, it can be said that

In, conclusion on the correlation between activity scoring and physiological stress response scoring: when the values of all subjects on both high and low load shifts are compared to the work activity, excellent correlations were found for all parameters indicative of activation of the SAM-axis. This should be seen as a reflection of work-induced mental activation. However, no significant correlations were seen with cortisol, an indicator of HPA-axis activation. This could generally be interpreted as a lack of correlation between the intensity or frequency of work activities and negative emotional experiences. In other words, the subjects did not find the increase in workload aversive. This conclusion supports the results of the subjective scoring comments. When analysing the low and high load results separately, no correlations were found by assessment according to activities and assessments according to physiological parameters. The fact that good correlations were found for the total number of subjects, but not for the subgroups can, with a fair amount of confidence, be ascribed to the fact that the experimental groups were too small. Indications repeatedly surfaced that telecommunications may be the one factor that is disliked when superimposed on the high load shifts

Part C: Preliminary indications of high allostatic loads

Although the determination of the indicators of accumulated long-term stress was beyond the brief for this work, a number of relevant observations were made. In normal populations salivary cortisol levels, as performed by the same standardised method used in this study vary between 4 and 10ng/ml at 08:00, and between 0.7 and 1.5ng/ml at 20:00hrs. The cortisol levels of the subjects in this study were generally much higher and varied between 4.0 and 16 ng/ml, with an average for all values, independent of work intensity, of 11.1 ± 4.1 ng/mL. The correlation between cortisol levels and BMI – indicators of acute as well as chronic, and chronic stress *per se* , was highly significant (r = 0.8277; p=0.0420) when values for the total group were analysed. This strong correlation as well as the high cortisol levels could reflect work stress, but could otherwise very well be a reflection of the life style of the workers. The mean blood pressure values of the total group were also above that of the normal range with systolic pressure of 136.7 ± 10.8 $mm.Hg^{-1}$, diastolic values of 95.1 ± 8.5 $mm.Hg^{-1}$, mean arterial pressures of 108.8 ± 8.5 $mm.Hg^{-1}$ and pulse pressures of 41.6 ± 8.3 $mm.Hg^{-1}$. The average heart rate for the group was 85.7 ± 8.8 $beats.min^{-1}$. The limited number of individuals on high and low work loads, respectively, as well as the fact that workers often alternate between high and low loads, unfortunately precludes the statistical comparison between high and low work loads. From the comparisons presented in this paragraph it is obvious that the physiological values of chronic stress parameters of the experimental group is above that of normal healthy population values. It is, however, highly likely that it reflects the total life style of the individuals rather than the work load and that such values would be found in many other work environments with life styles particular to the occupation. What was very obvious is the fact that low stress values, with regard to cardiovascular responses were consistently seen in the individual who gets regular exercise by walking to work.

Part D: Final conclusions and recommendations

The following conclusions can be reached

- ❑ That the work load between the presumed high load work station are indeed significantly different, not only in terms of trains and train orders, but also in terms of routine schedule notations, radio communications and telecommunications.
- ❑ The differences seen between workload activities and physiological responses are valid and not due to do to technical or design errors.
- ❑ The two parameters that were shown to be consistent indicators of acute workload increases and negative emotional experiences, respectively, were heart rate and cortisol. Blood pressure-related factors would, as reported in literature, seem to more reliant as a chronic stress indicator.
- ❑ The workload in terms of frequency of activities is reflected in the physiological stress response. However, this could be a reflection of the effect of allostatic load plus acute stress rather than merely acute stress as blood pressures are also increased.
- ❑ The increases in work load and psychological stress system activation is hardly ever accompanied by increases in a negative stress condition, i.e., the workers don't seem to resent high loads.
- ❑ The one factor, which would appear to cause negative stress and resentment, is telephonic communication when superimposed on a high intensity work schedule. It should be remembered that telecommunications are mostly the communication means when the normal routes fail or when special requests, beyond the work description, are made. Telecommunications would appear to be better tolerated at low workloads.
- ❑ When examining the possibility of the classification of stations into low and high stress areas it would appear if such a classification is indeed feasible as a good reflection of the physiological activation. This was seen in the activity score when the subjects as a whole were considered. As this impression was not supported by the results when the group was subdivided, the evaluation should be expanded to a larger number of subjects before absolute certainty can be reached.

- ❑ Not part of the brief of this study, but noticeable was the fact that these workers generally have allostatic loads, which, in the long run, could have health complications. This may, however, merely be the result of their life styles.

Recommendations:

It should be kept in mind that this was a pilot study with a limited number of experimental subjects and that the recommendations are based on the results thus obtained.

The major brief of this study was to physiologically assess the possibility of designing a formula, based on the activities of the train controlling officers, which could be used to assess the work load. The results of this pilot study showed that such a formula is feasible but that the following aspects should be addressed in creating the formula:

- ❑ **The activity-frequency of the different activities must be calculated separately and the appetitive-aversive quality must, where relevant, be a factor in subdividing the activities.**

- ❑ **Various activities must carry different weights in the formula.**
  This is supported by the fact that telecommunications, which often entails responsibilities beyond the job description of the train control officers, were found to be aversive when superimposed on a high workload. Indications were that telecommunications, as well radio communications, can become stressful when the first language is not the communication medium. The fact that radio communication can become stressful may very well be the result of the poor sound quality received by radio communication.

- **Assessment of the train control officers' perception of the stressfulness of the various activities is essential in the evaluation of the activities – and therefore in the compilations of the final formula. Ideally all control officers should complete a properly designed psychologically based score in which they rate the various activities in terms of stressor impact.** This recommendation is based on the fact that different activities have, on a physiological basis, already tentatively been shown to carry different stressor impacts**.** It was further noticed that what we would have thought to be aversive or non-aversive not always corresponded with the perceptions of the TCO's. The results of the psychologically based scores can, if necessary be verified physiologically, but the results should then be analysed against the allostatic load of the workers.

- **Personal aspects, such as ability and coping skills should be incorporate into the formula.** Although the aim is to evaluate the workload and not the worker such factors would become confounding aspects to the rest of the evaluations. Indications pointing towards this effect were borne out by results of this study. It is often good to have a system where the workload is specified at a certain competency or experience level.

- **The different shifts could very well carry different stress loads and should first be assessed by a subjective scoring system and, if needed, by heart rate, blood pressure and cortisol.** It is possible to assess the effect of continuous night shifts on the circadian secretory pattern of at least cortisol.

- **Allostatic loads**: It would only be fair, where relevant, to inform the workers about their high allostatic loads – especially their blood pressures – and to counsel them on the effect of life style on future health.

- **In conclusion it can be said that the results of the physiological-based pilot study indicated that a formula based on workload can, with certain prerequisites, give a fair reflection of the work stress. It is recommended that**

  a) The workload formula, based on the frequencies of the various activities as well as the results of operational research that considers the perceptions of the workers, be developed.

  b) The stations be evaluated in terms of the final formula and a distribution curve for each station be compiled

  c) At least 6 high stress and 6 low stress stations be identified and the validity of the activity-based scores be tested physiologically – in terms of a stat assessment of allostatic load and full shift heart rate, blood pressure and salivary cortisol – on each worker at the identified stations.

References

1 McEwen BS. Seminars in medicine of the Beth Israel Deaconess Medical center: protective and damaging effects of stress. N Engl J Med 1998;338:171-179.

2 Lovallo WR. Stress and health. Biological and psychological interactions. Sage Publications, California. 1997, p68

3 Kaplan JR. Primate models, cardiovascular disease. In: Encyclopedia of stress. George Fink. Academic press, San Diego, 2000, Vol 3, pp230-236.

4 Baker GJ, Suchday S, Kranz DS. Heart disease/attack. In: Encyclopedia of stress. George Fink. Academic press, San Diego, 2000, Vol 2, pp326-333.

5 Blessing WW. Regional blood flow. In: Encyclopedia of stress. George Fink. Academic press, San Diego, 2000, Vol 3, pp338-341.

6 Lovallo WR. Stress and health. Biological and psychological interactions. Sage Publications, California. 1997, p47.

7 Lovallo WR. Stress and health. Biological and psychological interactions. Sage Publications, California. 1997, p73

APPENDIX A

Perceived stress measure

ID: ……………………. **Date** ……/………/20…..
Station: ……………… **Time of day** ……….h………min **Shift time: ………………**

The following questions ask you about your feelings and thoughts DURING THE LAST MONTH. In each case, you will be asked to indicate how often you felt or thought a certain way. Although some of the questions are similar, there are differences between then and you should treat each one as a separate question.

		How Often In The Last Month? (circle your answer)				
		Never	**Almost Never**	**Some-times**	**Fairly Often**	**Very Often**
1	In the last month, how often have you been upset because of something that happened unexpectedly?	0	1	2	3	4
2	In the last month, how often have you felt that you were unable to control the important things in your life?	0	1	2	3	4
3	**In the last month, how often have you felt nervous and "stress"?**	0	1	2	3	4
4	In the last month, how often have you dealt successfully with irritating life hassles?	0	1	2	3	4
5	In the last month, how often have you felt that you were effectively coping with important changes that were occurring in your life?	0	1	2	3	4
6	In the last month, how often have you felt confident about your ability to handle your personal problems?	0	1	2	3	4
7	In the last month, how often have you felt that things were going your way?	0	1	2	3	4
8	In the last month, how often have you found that you could not cope with all the things that you had to do?	0	1	2	3	4
9	In the last month, how often have you been able to control irritation in your life?	0	1	2	3	4
10	In the last month, how often have you felt that you were on top of things?	0	1	2	3	4
11	In the last month, how often have you been angered because of things that happened that were outside of your control?	0	1	2	3	4
12	In the last month, how often have you found yourself thinking about things that you have to accomplish?	0	1	2	3	4
13	In the last month, how often have you been able to control the way you spend your time?	0	1	2	3	4
14	In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?	0	1	2	3	4

APPENDIX B

Maatstaf van die persepsie van stress

Perceived stress measure

ID: …………………….. Datum ……/………/20……
Stasie: ……………… Tyd van die dag……….h………min Skoftyd: ………………

Die volgende vrae handel oor u gevoelens en gedagtes GEDURENDE DIE AFGELOPE SKOF. In elke geval word u gevra hoe u gedink of gevoel het oor 'n besondere onderwerp. Alhoewel sekere van die vrae oënskynlik ooreenstem, verskil hulle tog en moet u asseblief elkeen as n afsonderlike vraag benader.

The following questions ask you about your feelings and thoughts DURING THE LAST SHIFT. In each case, you will be asked to indicate how often you felt or thought a certain way. Although some of the questions are similar, there are differences between then and you should treat each one as a separate question.

1	Hoe spanningsvol het u die skof gevind? 1 = Geen stres 5 = Baie stresvol How stressful did you experience the shift? 1 = No stress 5 = Very stressful	1	2	3	4	5
2	Was die skof meer of minder spanningsvol as gewoonlik? Was this shift more or less stressful than usual**?**	Meer spanningsvol Minder spanningsvol More stressful Less stressful				
3	**Was die skof vir u onaangenaam?** **Was the shift unpleasant / terrible / very bad?**	Ja Yes Nee No				
4	Was die skof vir u aangenaam / lekker? Was the shift pleasant / nice?	Ja Yes Nee No				

Rede waarom skof vir u aangenaam / onaangenaam was:
Reason why the shift was pleasant / unpleasant:

APPENDIX D

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
ESS ENTERPRISES
UNIVERSITY OF PRETORIA

P O BOX 14679 HATFIELD 0028
TEL: +27 (12) 420 4245
FAX: +27 (12) 362 5270
INTERNATIONAL: (+27-12)
e-mail: be@up.ac.za
website: www.be.up.co.za

REPORT

Validation of Spoornet Mental Workload Index Against an Allostatic Load Index

PREPARED BY

Prof M Viljoen
Dr N Claassen
Mr LT Hazelhurst
Department of Physiology
University of Pretoria

IN COLLABORATION WITH

Business Enterprises at University of Pretoria (Pty) Ltd
Graduate Center, First Floor, Entrance 2.27
Cnr. Lunnon and Herold Streets
Hatfield, 0083

November 2004

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 An Introduction to Workload as a Stressor

This work looks at the workload, and the environmental demands associated with it, as stressors that can lead to stress, or the stress response. The physiological parameters, which represent the peripheral expression of the stress condition in response to workload and demands, are investigated. In addition allostatic load is determined as a measure of the accumulated consequences of the stress impact on the body. Comparisons are made with the intensity of the occupational activity during the working period, and with the perceived stress of the working environment. The next couple of paragraphs provide a background to the investigation.

1.2 Meaning of the Word Stress

The meaning of the word "stress" has evolved from a) the early understanding of the word as referring to a stimulus, to b) stress seen as the response to a stimulus, to c) stress as a transaction. The stimulus model of stress saw stress as something that puts demands on the individual, for instance a heavy workload. This understanding of the word is generally considered to be wrong and the demands are now referred to as stressors rather than stress. The response model of stress refers to stress as the experiential and behavioural outcomes and includes the underlying physiological reactions. The response model thus sees stress as the reaction to a stimulus – a definition that includes acute reactions as well as the negative effects of chronic activation of the stress response. In terms of the response model, workload would be seen as the stressor and the physiological and psychological reactions as the stress or stress response. The transactional model of stress, also referred to as the process model, sees stress as a transaction between the individual and the environment and attempts to give a more holistic person-in context perspective (1). It should be obvious that the stress response is dependent on the nature of the stimulus and that the response may alter the characteristics of the stimulus, or the perception the individual has about the stimulus or stressor. However, the word stress is generally still used to refer to the response rather than to the transaction – while keeping in mind that the stress response is not static but a response that can change from moment to moment depending on the environmental demands and that the individual's reactions to the demands can influence the stressor value of the demand.

Many factors contribute to stress in the working environment. One of the major factors that determine stress in this environment is the workload. It is, however, important to remember that stress, as in this case caused by the workload, should not necessarily be seen as negative. In fact, stress underlies cognitive development, adaptation and the development of skills. It is now generally accepted that a controllable degree of stress is necessary for optimal performance. Too little stress as a result of a suboptimal workload may thus lead to a low performance or work output, while too much stress as a result of a heavy or uncontrollable workload may lead to a decline in performance level (2). One very important aspect that influences the effect of workload on both performance, and on health, is the perception of the individual. All stressors, including work, will necessary lead to a degree of psychobiological activation. The outcome is, however, strongly influenced by the individual's perception of the stressor. The negative feelings that the individual experience in the face of psychobiological activation, can be referred to as distress, while the high that individuals experience when the psychobiological activation is pleasurable is known as eustress (1,2). This would necessarily influence the work performance, as well as the degree to which the physiological stress response is activated.

1.3 Mediators of the Physiological Stress Response

The mechanisms through which a) stressful events and circumstances influence the homeostasis or internal stability of the mind and body, b) through which they influence performance, c) through which they impose wear and tear, and d) through which they can lead to disease, are multifactorial. Two of the main mechanisms through which these effects are mediated are known as the two main stress axes, i.e., the sympathodrenomedullary axis or SAM-axis and the hypothalamo-pituitary-adrenocortical axis or HPA-axis. Activation of the HPA-axis during stress leads to an increase in, amongst other cortisol, which in turn leads to increased access to energy stores from the conversion of other substances like lipids to glucose. Activation of the SAM-axis, which involves the activation of the sympathetic nervous system and release of catecholamines from the adrenal medulla, also helps to increase energy availability and, in addition, leads to an increase in cardiac output (through increased stroke volume and heart rate) an increase in blood pressure and a differential increase of blood flow and perfusion.

The stress response is supposed to be of limited duration and intended to help the individual to cope with demands. It is therefore adaptational in nature. The non-specific adaptive reactions are directed towards the mobilisation of the individual's reserves for energy and plastic modulation of the homeostatic responses and for a high level of functional activity (3). However, when demands on the individual become excessive the neuroendocrine activation associated with the stress response can become chronic with adverse effects on performance and health. Performance is, however, often maintained at the cost of health. The difference between the specific stress response and the non-specific stress response lies in the degree of specificity. Stressors that are very specific such as cold or hunger have their own specific stress responses that bring the internal homeostasis of the body back to normal through the normal negative feedback mechanisms of the body. However, any stressor that leads to psychobiological activation can give rise to the so-called non-specific stress response that involves, with minor variations, a relative non-specific neuroendocrine activation (4). This so-called non-specific stress condition or response can be see as a new homeostasis meant to enable the individual to cope, mentally and physically with the demands that initiated the response. Although the non-specific stress response involves virtually the whole body, two major neuroendocrine systems, as previously mentioned, are involved in the peripheral expression of the response, i.e., the sympathoadrenomedullary system and the hypothalamo-pituitary-adrenocortical axis.

In the following paragraphs we will shortly deal with the stress response in terms of the two main stress axis. It will also be shown that although we refer to the response as the "non-specific stress response" to distinguish it from small specific homeostatic deviations, with their individual correctional mechanisms, the non-specific stress response varies widely from individual-to-individual (interpersonal differences) and even within one individual at different times (intrapersonal). This makes it very difficult to be emphatic about the stress levels and stress responsivity of any individual. The discussion that follows will be limited to those physiological parameters which were possible to measure within the limitations set by the brief of the investigation, i.e., a non-intervention study within the working hours of the individuals.

1.3.1 The Sympathetic Nervous System and it's Relationship to the Rest of the Autonomic System

The sympathetic nervous system forms part of the autonomic nervous system, which consists of the sympathetic nervous system, the parasympathetic nervous system and the enteric nervous system. The sympathetic nervous system is so called because it is often concurrently activated or "in sympathy" activated during emotional and cognitive events. It is also associated with, and supports energy-requiring functions such as work. In this study three functions that is primarily associated with the sympathetic nervous system, i.e., heart rate, heart rate variability, and blood pressure are monitored. The second division of the autonomic system, i.e., the parasympathetic system, subserves energy-conserving functions and regulates energy intake such as feeding, digestion, absorption, as well as reproductive functions. It is more active during routine and vegetative functions, in contrast to the sympathetic system that is associated with action and stress. Although the main purpose of the study does not include assessment of parasympathetic function, an indication of its activation would be gained through heart rate variability analyses. Of more importance in terms of workload would be the balance between sympathetic and parasympathetic activation. The third division of the autonomic system is known as the enteric system. It is perhaps the only part of the autonomic system that is really autonomic and is involved exclusively with organs of digestion. The activity of the enteric division of the autonomic nervous system is partially modulated by the sympathetic and parasympathetic nervous systems (5).

The present study looks at the effects of work on the sympathetic nervous system as well as on autonomic balance - mainly in terms of sympathetic/parasympathetic activation.

1.3.2 The Sympathoadrenomedullary System

The sympathoadrenomedullary system, also referred to as the SAM-axis, consists of the sympathetic nervous system and the adrenal medulla. The sympathetic nervous system, as just discussed, comprises that part of the autonomic nervous system that is activated in the face of stressors, such as a heavy workload, in order to help coping with the demands on the person.

The effects of the sympathetic nervous system are mediated through the influence of its major neurotransmitter, i.e., noradrenaline (a catecholamine) on different kinds of adrenergic receptors. The adrenal medulla also secretes, amongst others, catecholamines such as noradrenaline and adrenaline and can in so-doing strengthen the effects of the sympathetic nervous system. When mass activation of the sympathoadrenomedullary system occurs it is often referred to as the fight-or-flight reaction. The overall effect is to aid the individual with coping with environmental and other demands mainly through an increase in perfusion of the vital organs with blood and an increase in blood sugar. These effects can, however become pathological in the chronic situation.

1.3.3 The Adrenal Medulla

The adrenal medulla is stimulated by sympathetic nervous system preganglionic fibres and can in a way be seen as an extension of the sympathetic nervous system. The adrenal medulla produces catecholamines, amongst others, adrenaline and noradrenaline, as well as other substances. This relatively simple view of adrenal medulla function is what is still being described in the majority of text books but over the last decade it has become clear that the adrenal medulla is not exclusively regulated by the sympathetic nervous system and that a host of other substances may modulate the influence of the sympathetic nervous system on medullary secretory activity and that differences in medullary function may in this way be accomplished, depending on the characteristic of the stressor (6). It is important to note that the secretory response of the adrenal medulla to stressors is dependent on the previous stress history of the individual. For instance, with exposure to the same stressor each day over an extended period of time subsequent exposure to the same stressor may induce a lower adrenal medullary response. This response could be seen as habituation. However, if the same individual which has been stressed by a particular stressor over an extended period of time is suddenly, in addition to the previous stressor, being exposed to yet another, i.e., novel, stressor the adrenal medullary secretory response may be significantly larger. This response would be known as sensitisation. The possible molecular mechanisms underlying habituation and sensitisation is largely beyond this writing but can be found in an overview by Kvetnansky (6). In terms of stress levels, or more specifically adrenal medullary stress response, of train control officers (TCOs) the implications from the above would be that individuals could adjust to the same workload when exposed to it over a period of time (habituation), but that introduction of significant changes could once again lead to higher stress levels than before (sensitisation). It should, however, be remembered that the workload and work environment are not the only potential stressors and that despite a degree of habituation the accumulative effects of long term above normal stress levels may adversely influence health – both mentally and physically. This will be addressed in more detail under allostatic load, a term that refers to the wear and tear of accumulative stress.

1.3.4 The Hypothalamo-Pituitary-Adrenocortical Axis (HPA-Axis)

In addition to the SAM-axis the other major stress axis is the hypothalamo-pituitary-adrenocortical axis or HPA-axis. This system is responsible for the glucocorticoid stress response. In humans cortisol is the major glucocorticoid. In the HPA-axis corticotropin-releasing hormone from the hypothalamic paraventricular nucleus, through it actions on the CRFR-1, stimulates the production of POMC with the subsequent release of adrenocorticotropic hormone or ACTH form the anterior pituitary gland. ACTH in turn stimulates the synthesis and release of cortisol from the zona fasciculate of the adrenal cortex. The H(hypothalamus)-P(pituitary)-A(adrenal cortex)-axis has a basal circadian rhythmicity or circadian rhythm, is sensitive to negative feedback by cortisol to the hippocampus, hypothalamus and pituitary gland, and is sensitive to stressors (7).

The circadian rhythm of cortisol secretion shows peak secretion in the early morning just before getting up and during early morning activity. It levels off later in the day but can increase as a result of meals or any of a variety of stressors. One of the major effects of stress including heavy workloads, unpleasant working environments or even continuous shift changes between day and night shifts, is, in fact, alterations of the circadian cortisol secretory pattern. Stress has several effects on the circadian rhythmicity of cortisol. The negative feedback of cortisol on the hippocampus, hypothalamus and pituitary (through which normal levels and changes in cortisol levels at different times of the day are regulated) appears to change under conditions of chronic stress and the sensitivity to feedback through cortisol seems to be reduced. One of the mechanisms may be through receptor downregulation with subsequent decreases in the sensitivity to cortisol (8). Acute stressors can cause acute increases in cortisol levels, which usually return to normal again within 2hrs. This is especially marked in the case of aversive stressors. Although acute stress can cause an increase in the basal-to-peak difference of cortisol the amplitude of the rhythm decreases in conditions of chronic stress as a result of the increase in baseline levels. As in the case of the SAM-axis, bouts of heterotypic acute stressor application during periods of chronic stress alter the response. It would seem that acute stress superimposed on chronic stress is dependent on the familiarity with the type of stressor. As in the case of the SAM-axis it appears that the response to an acute stressor in the chronically stressed would be less than expected if the stressor is homotypic (a repetition of what caused the chronic stress response). In contrast, if a heterotypic stressor is applied to a chronically stressed individual the response would be bigger (7). It is however possible that cross-tolerance to stressors may develop – especially stressors that involve the same neurological pathways (9). This once again can in theory be extrapolated to the working situation where high cortisol levels would then not be as significantly increased by increased levels of homeotypic stressors such as increases in the amount of work to which the individual is accustomed. In contrast, the response can be exacerbated when stressors other than the typical work stressors to which the individual is accustomed to are encountered. The implication that can be deduced from the previously mentioned fact (i.e., that the amplitude of the stressor-induced increase in cortisol levels, as well as the circadian amplitude difference to additional homeotypic stressors may be lower during chronic high stress) would be that those individuals with chronic high cortisol levels will not show the expected increase from basal trough values when stress levels increase and that this could be ascribed to the fact that the trough values are higher than normal.

There are many factors that influence the basal as well as the stimulated cortisol response to stress. One major factor is perinatal programming of the stress response that represents a major influence on the stress vulnerability of the individual and most probably can be held responsible for the largest part of the interindividual differences in the basal circadian levels of the individual (9). There are indications that gender is another contributor to interindividual differences and that distinct sexually dimorphic patterns of cortisol secretion may exist in the adult. There are indications that the cortisol levels are consistently higher in females than in males with further increases towards the middle of the menstrual cycle, just prior to ovulation (9). There are, however, some contradictions. Cortisol levels are further said to be increased in aging individuals with exaggerated responses to stress. This may, as in the case of chronic stress be the result of receptor down regulation. Such age-dependent increased cortisol levels and exaggerated responses to stressors can contribute to the development of many age and stress related disorders such as non-insulin dependent diabetes mellitus, osteoporosis, low immunocompetence and cognitive decline (9). Although this aspect may theoretically implicate that older workers could perhaps be less resistant to workplace stressors it is highly likely that the age-effect may be nullified by the fact that older workers may already be more familiar with the work, that they therefore perceive stressful events as less harassing, and that a degree of habituation may have developed.

1.4 Allostasis, Allostatic Systems, Allostatic Responses, Allostatic Load

Although the aforegoing part of this discussion was based on the principles of homeostasis, i.e., the relatively stable internal environment where deviations are corrected by negative feedback, it was also mentioned that in case of non-specific stressors the response may transiently give rise to a new homeostasis intended to help to cope with the stressful situation and that the homeostasis usually returns to normal when the stressful situation is over or the demand has been met. The latter situation implies flexible set points. The regulatory processes of the more flexible set point values around which the internal homeostasis can be regulated in the adaptive response, as seen during the transient development of a new homeostasis, and the regulation of the various functions around these new set points, should perhaps be referred to as homeodynamic rather than homeostatic regulation. There are however individuals that find even the homeodynamic concept too limiting and prefer to refer to allodynamic rather than homeostatic and homeodynamic regulation of the internal balance. In this categorisation it has become conventional to refer to the negative effects of the chronic stress response, as well as that of the cumulative effects of day-to-day activation of neuroendocrine systems as a result of intermittent heterotypic stressors, as the allostatic load. In this study we will be looking at the allostatic load as a measure of accumulated stress. It is thus necessary to shortly review the meaning of the different terms. The short overview that is to follow on allostasis, allostatic responses, allostatic systems and allostatic load is based on a number of publications (10,11,12,13,14,15,16) that are relatively freely available.

1.4.1 Allostasis

Allostasis can be seen as the ability to achieve stability through change. In terms of homeostasis one can compare this to the new homeostasis through which the mind and body attempt to cope with a stressor. In terms of coping with a heavy workload as the stressor, one can see it as the neuroendocrine changes that would help the individual to increase performance.

1.4.2 Allostatic Systems

The allostatic systems are those neuroendocrine systems that are activated in order to optimise coping and performance. In terms of homeostasis the allostatic systems can be equated with the stress systems.

1.4.3 Allostatic Response

The allostatic response could be seen as the stress response, in other words the response to stressors – the pattern of activation of the stress systems. The allostatic response varies from individual to individual and there are slight differences in the response to different stressors but the broad outline of the response is virtually the same as the broad outline for the non-specific stress response.

Various types of allostatic responses can occur. With the normal acute allostatic response, that is intended to be of benefit to the individual, the response which usually includes activation of the SAM-axis and the HPA-axis with subsequent increases in heart rate, blood pressure, blood sugar and other physiological changes, the response is shut off after the stressful event and the physiology of the body will return to normal. As long as the allostatic response is limited to the period of the work, or the demands on the individual are not excessive, protection via adaptation predominates over adverse consequences. However, exposure over extended periods of time to the allostatic response can have pathophysiological, and sometimes also psychopathological, consequences. Chronic stressful situations and other factors may lead to abnormal allostatic responses which can, in turn, give rise to a high allostatic load. The concept of allostatic load will briefly be discussed after dealing with the various types of allostatic responses.

Examples of abnormal allostatic responses that may lead to abnormal allostatic loads include a) chronic activation of the stress or allostatic systems as a result of frequent different stressful situations (heterotypic stressors) where there is hardly any time for the stress systems to operate at baseline values, b) chronic activation of the stress systems as a result of non-habituation of the individual to homeotypic stressors where once again the stress systems continuously operate at high levels, c) the inability of stress systems to appropriately shut down once they have been activated, and d) the inability of the body to mount the necessary stress response in the face of a stressor. Although the latter is often dependent on early childhood development of the stress systems, pathological hyporesponsivity is also found with long-term subordinance and the feeling of hopelessness that accompanies it. It is said that the type of hyporesponsiveness where low reactivity and basal activity of the HPA-axis are found could be a characteristic of the chronic fatigue syndrome – this is however still controversial. In the first three situations the body would chronically be subjected to high levels of, amongst others, activation of the SAM-axis and HPA axis with physiological disturbances such as increases in heart rate, blood pressure, cortisol levels and abnormal or low heart rate variability. The potential end-result of this chronic activation includes, depending on other factors such as genetics, feelings of constant fatigue and perhaps demoralization and hostility, the inability to concentrate, irritability, depression, hypertension or in those so inclined hypotension, non-insulin dependent diabetes mellitus, deposition of abdominal fat, osteoporosis, increased risk for cardiovascular incidences such as myocardial infarcts, artherosclerosis and other cardiovascular pathology, as well as an increased vulnerability to infection and cancer due to a suppression of immune function. There are indications that allostatic responses can become more intense with aging and it would take longer for the body to recover to baseline value. However, one should perhaps ask whether this is the effect of chronological aging *per se* or not merely a result of the general decline over long-term exposure to allostatic responses. Although there are indications that females are perhaps more stress responsive than males, there are certain factors that protect females against the negative effects of high stress system activation. Oestrogen, for instance is said to protect the cardiovascular system of females against the effects of high activation of stress systems, but this effect is naturally lost in postmenopausal women. It would further appear that a decline in oestrogen levels might contribute to cognitive decline as a result of increases in the activity of the HPA-axis.

1.4.4 Allostatic Load

The allostatic load can in a sense be seen as the same as the effect of the long-term or chronic so-called non-specific stress response. It is the price you pay for the repeated or chronic activation of the allostatic response – a response which in the short term is intended to be advantageous. Another way of putting it would be to define it as the wear and tear that results from chronic overactivity or underactivity of the allostatic systems. The allostatic load is the product of the accumulated stressors over an extended period. This includes the effects of minor day-to day stressors and the more dramatic or traumatic events. There can be no doubt that the type of work, the working environment and the working hours could be seen as major determinants. In this the perception of the individual plays a significant role. It is one of the major factors that determine whether he or she will be sensitised or will habituate to a certain stressor, i.e., how he or she will cope with the different day-to-day stressors and for how long the neuroendocrine systems will remain activated. Other factors which can be seen as instrumental in modulating the characteristics of the stress or allostatic response are the general physical health and the life style of the individual – including exercise, the diet, alcohol intake and smoking.

Some psychosocial factors that may influence allostatic responses and the allostatic load have previously been touched upon in this writing. Other psychosocial factors that may have a bearing on this study include observations that the effects of increased allostatic load such as artherosclerosis are more prevalent in socially dominant males of unstable social hierarchies and in subordinate females, observations that low job control predicts an increased risk for coronary heart disease and that high job strain can lead to a chronic increase in ambulatory blood pressure, left ventricular mass index and the progression of artherosclerosis.

In summary, to quote, Bruce S McEwen, Rockefeller University, Bethesda, 2000 (17):

allostasis *The ability to achieve stability, or homeostasis, through change, as defined by Sterling and Eyer, is critical to survival of an organism by promoting adaptation, or the reestablishment of homeostasis, to an environmental challenge. Through allostasis, the autonomic nervous system, the hypothalamic-pituitary-adrenal (HPA) axis, and cardiovascular, metabolic, and immune systems protect the body by responding to internal* *and external challenges. Allostasis is achieved through the action of mediators, such as the catecholamines and glucocorticoids. Allostatic load is the price of this accommodation and constitutes the wear and tear on the body from the chronic overactivity or underactivity of allostatic systems.*

allostatic load *The excessive level, over weeks, months and years, of mediators of allostasis, resulting either from too much release of these mediators or from the inefficient operation of the allostatic systems that produce the mediators and fail to shut off their release when not needed.*

McEwen 2000 (17)

1.5 Activities of the Two Main Stress Axes as Indicators of Workload

1.5.1 Heart Rate

Heart rate could very well be one of the most general mechanisms through which the stress response is assessed. From a research point of view heart rate can give an indication of sympathetic activation or autonomic balance. Heart rate can be determined by a variety of methods ranging from counting the pulse by palpitation over an artery, to counting the QRS waves per unit time on an electrocardiogram (ECG), to electronic monitors designed specifically for heart rate determinations. The palpitation-dependent determinations of heart rate is not an option for research purposes as it requires the constant presence of another individual and a possible influence on baseline and other responses. It also prevents the individual being tested from doing any work. With ECG determinations of heart rate the QRS waves per unit time are counted or the time between the waves (interbeat intervals) measured and the rate calculated from that. An interval between heart beats of 1000 msec would, for instance, give a heart rate of 60 beats per minute, that is 60 times the 1000 msec intervals which would equal 60 000 msec or one minute. In measuring the beat-to-beat interval one would perhaps expect to measure the P-wave-to-P-wave interval as the P-wave is initiated from the sinus node region that initiate cardiac contraction. The major reason for measuring the QRS intervals is that the QRS is easier to detect whereas the p-wave, which initiates the cardiac activation, is smaller. The p-wave starts form the sinus node of the heart and its discharge depends on the depolarization of neural tissue, modulated by the accelerating influences of the sympathetic nervous system and the slowing influences of the parasympathetic nervous system. Monitoring of the beat-to-beat changes in heart rate intervals could therefore give an indication of the sympathetic/parasympathetic balance and of the activity of the SAM-axis (18).

From a stress perspective it is necessary to note that the sympathetic nervous system, and by implication the adrenal medulla, is activated by most stressors and that the sympatho-adrenomedullary system, in turn, does not only stimulate heart rate through its effects on the sinus node, but can also influence the contractility of the cardiac muscle and thereby the systolic blood pressure. Acute reactions to stressors can be judged by observing the increase from baseline in heart rate. What is interesting is the fact that heart rate increases, despite the fact that significant transient increases may occur upon stressor application, do not seem to predict disease onset and cause. It would appear that increases in blood pressure could be more predictive in nature (18). It is suggested that the parasympathetic (vagal) influence which slows the heart down and which is the primary regulatory factor during resting conditions, may be a contributing factor to this phenomenon.

Although heart rate can serve as an index of stress, especially of short-term stress, it is necessary to note that many factors can influence the recordings, amongst others physical activity. This makes it problematic to distinguish between the emotional-cognitive aspects and the more physical aspects. In the present study the volunteers were not exposed to overt physical activity, as their job is more of a cognitive nature with some emotional stress factors superimposed by problems in the controlling of trains. It is, however, true that factors which do form part of the execution of their work can also influence the heart rate, including minor forms of motor activity such as writing, reaching for instrumentation and moving between points, as well as changes in respiratory patterns such as breath holding, transient shifts in attention and muscle tension (18). Luckily many of these phasic changes disappear after a while and the heart rate recordings that were performed in this study were of a continuous nature. Should continuous increases in muscle tension be present it could have influenced the heart rate, but it is to remember that increased muscle tension is yet another expression of stress.

1.5.2 Heart Rate Variability

Heart rate variability (HRV), which is based on small changes in the beat-to-beat QRS intervals (also referred to as R-R intervals), provides a non-invasive indication of autonomic function. It is generally thought that increased sympathetic stimulation decreases heart rate variability and that increased parasympathetic stimulation increases heart rate variability (19). Of perhaps greater importance is the implication of this, i.e., that the degree of variability may be indicative of cardiac health. Decreased variability is, for instance, being associated with increased mortality after myocardial infarct. This would make sense in view of the effects of autonomic influences on heart rate variability and the fact that heart failure is characterized by increased sympathetic and decreased parasympathetic nervous system activity (19). It would therefore be valid to suggest that not only would heart rate variability determinations be a good tool for the assessment of stress-induced sympathetic activation, but also that a decrease in baseline heart rate variability could also be a measure of the allostatic load.

As instantaneous changes in beat-to-beat intervals often become obscured during analysis by means of standard deviations, histograms and spectral analyses of HR-recordings, specialized techniques such as Poincaré plotting, can be employed to assess HRV and autonomic nervous system shifts.

1.5.3 Blood Pressure

Blood pressure is the outward force exerted against the walls of the blood vessels. When we talk about blood pressure we usually refer to the pressure in the arteries. When we talk about the pressure in the veins we specifically talk about venous pressure. Four pressures are of important for the purpose of this study, i.e., systolic pressure, diastolic pressure, pulse pressure and mean arterial pressure. The blood pressure varies depending on whether the heart is in contraction or relaxation and we thus refer to systolic pressure (when the heart contracts), or diastolic pressure (when the heart is in relaxation). The mean systolic pressure, as conventionally measured, is about 120mmHg in a young adult male and the diastolic pressure about 80mmHg. The term pulse pressure is the amount the pressure increase from diastole to systole. The formula for the calculation is thus: pulse pressure = (systolic pressure – diastolic pressure), or 120 – 80 = 40mmHg. It is obvious that the pulse rate is the same as the heart rate (20). The mean blood pressure or rather the mean arterial pressure is the average pressure throughout the cardiac cycle or, put in a different way, the average pressure to which the walls of the blood vessels are being subjected. As the period of each contraction of the heart is slightly shorter than the associated period of relaxation, the mean arterial pressure or MAP is slightly lower than the sum of the two divided by 2. The mean arterial pressure can be calculated by the formula: mean arterial pressure = diastolic pressure + ((systolic – diastolic)/3)

Blood pressure is the result of the pressure of the blood against the blood vessels which offers resistance to expansion of their diameter. The volume of the blood being pumped out of the heart, that is the activity of the heart, as well as the degree of resistance offered by the blood vessels, i.e., the peripheral resistance, determine the blood pressure. Both cardiac output and peripheral resistance are influenced by stress hormones and the autonomic nervous system (20). The effects of stress on blood pressure will be returned in a later paragraph.

Blood pressure can be measured in different ways. Two commonly used techniques, the auscultatory and oscillometric techniques employ an occlusion cuff through which pressure is applied over the artery to first interrupt and then progressively restore blood flow through the artery. A number of research-orientated techniques, including cuff-tracking, vascular unloading, arterial tonometry and pulse transit time provide methods to determine blood pressure beat-by-beat, non-invasively, each with its own advantages and disadvantages. Blood pressure in humans are routinely, for clinical purposes, measured by the auscultatory method by means of an instrument known as the sphygmomanometer where an inflatable rubber cuff that is connected to a mercury-containing pressure meter is place round the upper arm. Blood pressure varies depending on where it is taken but is most commonly taken over the brachial artery. To measure the systolic and diastolic pressures a stethoscope is placed on the brachial artery at the elbow fold and the cuff inflated to above the expected systolic pressure. At this stage blood flow through the brachial artery is stopped during cardiac relaxation as well as during cardiac contraction – due to the occlusion by the inflated cuff. The cuff is then slowly deflated while listening with the stethoscope over the brachial artery. At the point that the systolic pressure becomes equal to the cuff pressure blood will start flowing through the artery during systole but not during diastole and this intermittent flow can be heard as a tapping sound. The pressure reading on the sphygmomanometer now represents the systolic pressure. As the cuff is deflated further the pressure in the cuff eventually becomes lower than the diastolic pressure and blood can now also flow through the brachial artery during diastole, i.e., blood flow is now continuous and the tapping sound disappears. This pressure is taken as the diastolic pressure (21).

Stress can lead to high blood pressure or hypertension. Although not all cases of hypertension can be associated with stress, and the cause of hypertension is no doubt multifactorial, the evidence that stress can exacerbate the influence of most other factors involved in an increased blood pressure is overwhelming (21). The contribution of stress to high blood pressure is largely mediated through the influence of the sympathetic nervous system. It has now become clear that there are high and low responders and that people whose blood pressure increase significantly in the face of stress could be more prone to the development of hypertension. According to the diathesis-stress model of essential hypertension (22), hyperreactivity to stress is said to lead to hypertension only in the presence of other predisposing negative psychosocial factors, and individual differences in personality or other stable behavioural traits should be seen as mediators of the reactivity-psychosocial hypertension relationship (22). This assumption is supported by finding showing that sympathetic reactivity could very-well be a function of interactions between job stress, other environmental demands, family history of hypertension, ethnicity and emotional disposition. Family history of hypertension includes a possible genetic contribution but also dietary aspects like sodium intake, and the family psychosocial environment. With reference to ethnicity, there are indications that individuals of non-Caucasian origin may show a greater vascular reactivity to challenges, may excrete less sodium relative to intake, and other signs that they may be at greater risk for hypertension (23,24,25).

With reference to the effect of acute and chronic stress in the work place it should be remembered that it is not only the work load that influences blood pressure but that emotions like anxiety, negative moods, depression and lack of social support can also be risk factors for the development of hypertension (21,23). The increase in blood pressure induced by the work environment may be carried over and blood pressure would not necessary returned to normal when leaving the premises. This would surely reflect as an increase in allostatic load.

1.5.4 Cortisol

The HPA-axis as one of the two major stress axes was briefly discussed in an earlier section. In this study the levels of cortisol were measured as an indication of the activation of the HPA-axis. In the previous paragraphs we briefly looked at the effect of stress on heart rate, heart rate variability and blood pressure and although the HPA-axis does have an influence on these factors the influence of the sympathetic nervous system is much more significant. Before any further discussion on the cortisol response it should be mentioned that, in contrast to the cortisol response, the reactivity of the sympathetic nervous system to stress would appear to be less dependent on the nature of the stressor. Although cortisol levels can to a degree increase in the face of different kinds of stress, high cortisol- reactivity appears to be associated with aversive situations, the anticipation of aversive events, subjective states of fear, frustration, negative emotions accompanying failure, effort without positive outcomes, situations with potential negative evaluations of the self, loss of status amongst equals, low controllability and similar situations. (26).

A wide range of individual differences can be found with regard to cortisol reactivity. In younger subjects (20-29) women seem to have lower 24hr plasma cortisol secretion as well as lower morning cortisol peaks, but their age-related increases in basal output would appear to be more significant than men. In contrast to some reports about a higher stress responsitivity in women, other publications also showed cortisol stress responsiveness, similar to sympathetic responsiveness, that appear to be lower in women than in men – indicating perhaps that men are more responsive to threatening or challenging cues. (26).

Cortisol can be measured in blood, urine or saliva. In human research the collection of blood samples may, however, be problematic as venipuncture in itself is experienced as a stressful event by many and the taking of blood samples, especially serial blood samples would become a confounding factor. In the present study cortisol was determined in saliva. Saliva is considered an unobtrusive specimen source for cortisol – especially for studies performed outside the laboratory situation. Salivary cortisol represents the unbound fraction of cortisol, i.e., the biological active part and the correlation between salivary cortisol and the unbound fraction of cortisol extracted from serum or plasma is very good.

1.6 Stress Reactivity in General

Early on in this writing it was referred to the fact that the complex neuroendocrine activation in the face of stress is often referred to as the non-specific stress response merely to differentiate between this complex stress response and the specific homeostatic mechanisms that operate in the face of specific disturbances for which specific feedback mechanisms (specific stress responses) exist. It was also shown that individual stress reactivity and general health often form the basis of the eventual outcome of stressful events or periods. It is perhaps necessary to briefly touch upon some aspects, not discussed before, which influence stress reactivity.

1.6.1 Differential Response Patterns

It can for instance now be taken for granted that differential responses of the HPA-axis and the SAM-axis, i.e., differential metabolic and cardiovascular stress responses, can occur depending on access to and activation of appropriate coping mechanisms. Two typical responses can be distinguished, i.e., the activational pattern, also referred to as the defense pattern and the inhibitory or vigilance pattern. There is also a pattern referred to as the defeat pattern where the individual feels totally out of control and gives up on any form of coping with dire physical and psychological consequences, but this is not to be discussed here. It is said that the type-1 pattern (the activational pattern) is marked by increased skeletal muscle vasodilation, increased cardiac output, increased systolic blood pressure and beta-1-adrenergic tone, while type-2 (the inhibitional pattern) response is marked by skeletal muscle vasoconstriction, increased diastolic blood pressure, increased total peripheral resistance and increased alpha-adrenergic tone (23). Although increases in blood pressure during stress are generally of a type-1 or type-2 pattern, a mixed response pattern may also occur which appears to be dependent on the demands of the task itself, its context, differences in the person's history, his perceptions, response styles and other factors (23,27). Of interest is the fact that these reactions are very much dependent on the personality type and that many of the differential blood pressure responses are expressed when individuals are harassed in the work place.

1.6.2 Anticipation as Stressor

Emotions like anticipation, anxiety and fear can also drive the stress response. Such negative emotions can act as acute stressors and often have adaptational and coping value, but chronic anticipatory anxiety can, because it drives the stress response, lead to chronic high levels of stress hormones and sympathetic nervous system activation. The HPA axis and the amygdala are two of the major systems involved in the development of anticipatory angst (28).

It is known that this type of anticipation may contribute to the chronicity of stress response activation. The measurement of the stress or allostatic response to a stressful situation is often confounded by the fact that individuals would arrive at a situation already stressed in anticipation of the stressful situation. In theory this can, in the short-term, be seen as a reactions that prepare the individual for the job to come or the stressor about to be faced. If the period of anticipation is prolonged, such as in the case where a worker finds his work extremely stressful and starts worrying about it long before his day or shift starts it will eventually contribute to a chronic high allostatic load and the associated health consequences The anticipatory response makes it extremely difficult to assess the impact of job strain on the individual but it can with a fair degree of confidence be assumed that the anticipation will be relatively low in low stress working environments and high in high stress working environments (17,28).

1.6.3 The Orienting Reflex

In addition to anticipation, another factor that may influence the stress level determinations at the onset of a work shift is the orienting reflex. This reflex was first described by Pavlov in 1927 (29) who preferred to refer to it as the “investigatory” or “what-is-it” reflex. Sokolov (30), who applied a cognitive approach to the reflex, initiated the idea that the orienting reflex has a comparator component, comparing present observations or input to past representations or past input. The orienting reflex would especially be elicited when there is a discrepancy between current and past or stored input – that is, when the situation is new or has changed from the previous occasion or presentation (30). The equivalent in the work environment is easily imagined. The orienting reflex is associated with an increase in sympathetic activity. Although it is often measured by the skin conductance response – neurally it reflects exclusive sympathetic control of palmar sweat glands (31) - other techniques of sympathetic nervous system activity can also be used. There may be habituation of the orienting response and it can surely be expected that novices will show a stronger orienting reflex under the same conditions where a more experienced worker may show more habituation.

1.7 Aim of the Study

The primary aim of the study was to validate the developed MWL-index against certain parameters of stress.

Secondary aims were to:

1) examine the physiological changes in terms of the two main stress axes as measures of the work stress at different train controlling venues in order to assess workload or work stress at the different stations,
2) assess measures of the allostatic load in an attempt to validate the wear and tear of accumulated stress on the individuals – another approach to work stress at the different venues and
3) compare the MWL-index to the subjective perceptions of the observer and ratings based purely on time line analysis.

2. Materials and Methods

2.1 Subjects

Twenty train control officers from seventeen stations across South Africa took part in this study. Anthropometric measurements of participants are listed in Table 1.

2.2 Experimental Trials

Testing commenced at the start of the train control officers shift at 06h00. In four of the cases the officers had reached a ‘gentlemen’s agreement’ and only started at 07h00. The results of these subjects then only begins at 07h00. The subjects performed their handing over shift duties and were then informed about the nature of the study. Once the subjects were set the experimental procedures and aims were explained to them and testing commenced.

2.3 Measurements

2.3.1 Cohen's Measure of Perceived stress

Subjects completed the Cohen's perceived stress scale to assess their non-specific, appraised stress during the last month. This questionnaire consists of 14 items of which 7 are positively formulation (e.g. "In the last month, how often have you felt that things go your way?") and 7 negatively formulated (e.g. "In the last month, how often have you felt that you were unable to control the important things in your life?").

2.3.2 Blood Pressure

Blood pressure i.e. diastolic and systolic blood pressure was measured with a digital electronic blood pressure meter (ALP K2, model DS-125D, Japan). Three consecutive blood pressure measurements were taken at the commencement of the shift and every two hours thereafter during the shift till 14h00. The average of the three measurements is reported.

2.3.3 Heart Rate Recording and Analysis

A Direct Wired POLAR® (Mini Mitter Co, Inc. Bend, OR USA) heart rate belt was attached to the subject. The electrodes were first primed with distilled water. Heart rate data is transferred instantaneously via a direct wire to the Mini-Logger® Series 2000 (Mini Mitter Co, Inc. Bend, OR USA) and stored until download. Inter-beat interval is sampled by timing the number of milliseconds between triggering's of the Polar heart rate transmitter. The number of milliseconds between adjacent pulses from the transmitter are counted by the logger and recorded.

Heart rate data recorded continuously during the 8 hour exposure was then downloaded to a laptop computer using the Mini-Log^{TM} 2000W for Windows® (Mini Mitter Co, Inc. Bend, OR USA). The binary file created was then converted to an ASCII file where it was edited in Microsoft Excel in order for heart rate variability analysis. The text file created was imported into HRV analysis software (biomedical Signal Analysis Group, Finland). The program generates a report sheet and exports results as a text file.

2.3.4 Cortisol Collection and Analysis

Cortisol is found in the blood bound to protein or free (3 – 5 %). It is common to measure cortisol in either blood (where total cortisol is measured) or in saliva where the free fraction is measured. The rhythm or cortisol secretion corresponds to the sleep-wake cycle rather than the day-light cycle of melatonin secretion. Cortisol is a steroid hormone and therefore relatively stable.

Saliva was collected in a clean collection test tube, at the start of the shift and every two hours thereafter till 14h00, and then stored on ice for centrifugation. The collected saliva was centrifuged at 3300 rpm for 30 minutes. An aliquot of the supernatant was transferred into two 1.5 Ml eppendorf tubes and stored at –70 °C until analysis.

Cortisol concentration was determined by ELISA using a DX-SLV-2930 cortisol saliva kit (AEC Amersham Pty, Ltd., South Africa).

2.3.4.1 Principles of the Test

The solid phase enzyme immuno-assay for cortisol is a competitive type immuno-assay wherein horseradish peroxidase-labeled cortisol (HRP-cortisol) competes with cortisol present in the subject sample for a fixed and limited number of antibody sites immobilised on the wells of the microstrips.

Once the competitive immunoreaction has occurred, the wells are rinsed, and the HRP-cortisol fraction bound to the antibody in the solid phase is measured by adding a chromogen/substrate solution that is converted to a blue compound. After 15 minutes of incubation, the enzymatic reaction is stopped with sulphuric acid that also changes the solution to a yellow colour. The absorbance of the solution, photometrically measured at 450 nm, is inversely related to the concentration of cortisol present in the sample. Calculation of cortisol content in the sample is made by reference to a calibration curve.

Expected values range from 0.4 – 1.0 µg / dL (11 – 28 nl / L) if the samples were taken at 08h00.

2.4 Estimation of Mental Workload

Time line analyses, as indicator of mental workload, were recorded during the 8 hour testing periods. The same observer was responsible for the time line analysis of all the subjects tested. The main activities recorded were updating of train diagrams (scheduling), radio and telephone communications, as well as the number of trains and authorisations during the 8 hour testing period.

2.5 Statistics

All values are represented as mean ± standard deviation. A Two Sample T-test was used to determine if differences exist in terms of allostatic load variables between the identified Low and High stress groups respectively. A two-way repeated measures Analysis of Variance (AOV) was used to evaluate between subject groups (Low and High stress) differences and within subject differences (responses over an 8 hour work shift). Group*Time interaction was also evaluated to determine whether both groups respond in a similar fashion over the 8 hour work shift. If between and/or within subject analysis indicates any significant differences, but Group*Time interaction exists, then the two groups respond differently to the stressor over time and care must be taken in the interpretation of the main effects e.g. groups and time. Statistix Ver. 8 program was used for the statistical analysis. Prof PJ Becker (MRC) consulted on the statistical analyses.

3. RESULTS

Twenty train control officers at various venues were selected by Spoornet to take part in the physiological validation of MWL-index by assessment of stress levels in the train control officers stationed at various venues. Selection was performed on the basis of an estimation of the stressor impact at the different venues as calculated by the MWL-index. Although the investigators (researchers involved in the present study) new that the workloads at the different venues differed, i.e., ten were high workload stations and ten low workload stations, the investigators did not have insight into which venues were high stress and which low stress until completion of the practical and statistical analyses of their own results.

The results are being presented in the following order:

1. Anthropometric data and absolute recorded values (3.1)
2. Comparisons between Spoornet models of workload at the different venues and the physiological indicators of individuals working at the respective venues (3.2)
3. Physiological parameters: Separation of work stations into high and low stress as extrapolated from the stress levels of the individuals at the work stations (3.3)
4. Integration of Spoornet model and combinations of experimental test parameters and online time analysis (3.4)

3.1 Anthropometric Data and Absolute Values of Physiological Stress Indicators

The experimental group consisted of seventeen men and three women. Other anthropometric details of the twenty experimental persons are presented in Table 1.1.

Table 1.1: Anthropometric data of TCOs selected by Spoornet to participate in the validation of the MWL-index.

Subject	Gender	Age (y)	Mass (kg)	Length (m)	BMI (kg/m^2)	SA (m^2)
1	Male	32	94.0	1.841	27.7	2.20
2	Female	23	50.8	1.565	20.7	1.50
3	Male	31	84.1	1.738	27.8	2.03
4	Male	48	120.4	1.731	40.2	2.44
5	Male	34	119.5	1.820	36.1	2.48
6	Male	44	88.6	1.776	28.1	2.11
7	Male	45	76.8	1.643	28.5	1.89
8	Male	49	88.4	1.763	28.4	2.10
9	Male	41	84.5	1.672	30.2	2.00
10	Male	38	82.6	1.703	28.5	2.00
11	Male	43	145.0	1.760	46.8	2.70
12	Male	31	119.3	1.870	34.1	2.51
13	Male	52	86.5	1.781	27.3	2.08
14	Male	42	81.0	1.737	26.8	1.99
15	Male	59	113.6	1.799	35.1	2.41
16	Male	54	94.1	1.743	31.0	2.16
17	Male	35	95.9	1.840	28.3	2.23
18	Male	46	71.2	1.635	26.6	1.82
19	Female	28	48.7	1.504	21.5	1.44
20	Female	29	59.2	1.635	22.1	1.65
Mean		40.2	90.2	1.728	29.8	2.09
SD		9.7	24.3	0.095	6.2	0.33
Min		23	48.7	1.504	20.7	1.44
Max		59	145.0	1.870	46.8	2.70

BMI = Body Mass Index (kg / m^2); SA = Surface Area (Mass$^{0.425}$ x Height$^{0.725}$ x 71.84)

The salivary cortisol levels at 06:00, 08:00, 10:00, 12:00 and 14:00, the mean and standard deviation of cortisol levels for each train controlling officer from 06:00 to 14:00, the mean and standard deviation of cortisol for each from 08:00 to 14:00 and the minimum and maximum values for each individual, as well as the means, standard deviations, minimums and maximums for the total group at each of the five time intervals respectively, can be seen in Table 1.2. It is graphically represented in Figure 1.1.

Table 1.2: Cortisol levels of TCOs over an eight hour shift at the workstations selected by Spoornet for the MWL-index validation. Mean and standard deviation of cortisol values for each TCO was calculated over the whole eight hour shift and from 08:00-14:00. Min and maximum values for each TCO are also reported. Mean, standard deviation, minimum and maximum values for every 2 hours from 06:00 – 14:00 are also reported for all the TCOs.

	Cortisol (ng/ml)								
Subject	06:00	08:00	10:00	12:00	14:00	Indv. Mean 06:00-14:00	Indv. Mean 08:00-14:00	Min.	Max.
1	5.5	3.0	5.0	2.5	4.0	4.0 ± 1.3	3.6 ± 1.1	2.5	5.5
2	6.0	3.0	3.0	2.0	2.0	3.2 ± 1.6	2.5 ± 0.6	2.0	6.0
3	3.5	3.5	2.0	3.5	4.0	3.3 ± 0.8	3.3 ± 0.9	2.0	4.0
4	5.5	3.5	3.5	4.0	2.0	3.7 ± 1.3	3.3 ± 0.9	2.0	5.5
5	4.5	5.5	3.5	1.0	2.0	3.3 ± 1.8	3.0 ± 2.0	1.0	5.5
6	7.5	5.0	4.0	3.0	1.5	4.2 ± 2.3	3.4 ± 1.5	1.5	7.5
7	6.0	3.5	2.0	8.0	2.5	4.4 ± 2.5	4.0 ± 2.7	2.0	8.0
8	11.0	7.5	5.0	4.0	3.5	6.2 ± 3.1	5.0 ± 1.8	3.5	11.0
9	8.0	4.5	4.0	2.0	2.5	4.2 ± 2.4	3.3 ± 1.2	2.0	8.0
10	9.0	3.5	6.5	5.0	5.0	5.8 ± 2.1	5.0 ± 1.2	3.5	9.0
11	12.0	8.0	6.5	7.5	6.5	8.1 ± 2.3	7.1 ± 0.8	6.5	12.0
12	6.5	2.0	1.5	2.5	1.5	2.8 ± 2.1	1.9 ± 0.5	1.5	6.5
13		5.0	3.5	5.0	5.5	4.8 ± 0.9	4.8 ± 0.9	3.5	5.5
14	10.0	5.0	8.0	2.0	6.0	6.2 ± 3.0	5.3 ± 2.5	2.0	10.0
15	17.5	9.0	6.0	4.0	4.5	8.2 ± 5.6	5.9 ± 2.3	4.0	17.5
16	9.0	8.5	7.5	4.5	4.0	6.7 ± 2.3	6.1 ± 2.2	4.0	9.0
17	6.5	3.5	3.0	3.0	2.0	3.6 ± 1.7	2.9 ± 0.6	2.0	6.5
18	5.0	2.5	3.0	3.5	3.5	3.5 ± 0.9	3.1 ± 0.5	2.5	5.0
19		3.0	3.0	5.0	3.0	3.5 ± 1.0	3.5 ± 1.0	3.0	5.0
20		3.5	2.5	2.5	4.0	3.1 ± 0.8	3.1 ± 0.8	2.5	4.0
Mean	7.8	4.6	4.2	3.7	3.5				
SD	3.4	2.1	1.9	1.8	1.5				
Min	3.5	2.0	1.5	1.0	1.5				
Max	17.5	9.0	8.0	8.0	6.5				

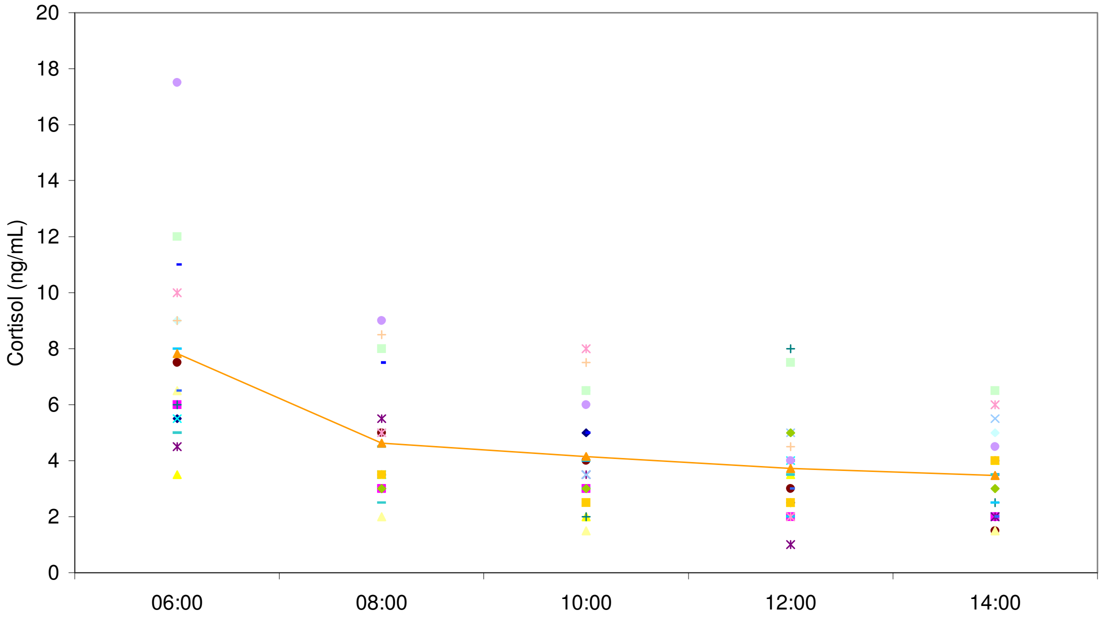


Figure 1.1: Cortisol levels for each individual from 06:00 to 14:00.

Table 1.3 presents the results of Cohen's perceived stress measurement

The means, standard deviations, as well as minimum and maximum values for blood pressure were calculated in the same way as that for cortisol. The systolic blood pressure can be seen in Table 1.4, a graphic presentation of the systolic blood pressure in Figure 1.2, the diastolic blood pressure in Table 1.5, a graphic presentation of the diastolic pressure in Figure 1.3, the mean arterial blood pressure in Table 1.6, a graphic presentation of it in Figure 1.4, the mean pulse pressure in Table 1.7, a graphic presentation of it in Figure 1.5, the heart rate values as recorded by Minimitter in Table 1.8, the graphic presentation of it in Figure 1.6, and a graphic presentation of the heart rates as determined by automatic blood pressure monitor in Figure 1.7

Table 1.3: Cohen's perceived stress measure completed by TCOs at the beginning of their shift on the day of validation.

Subject	Cohens perceived stress measure	Associated level of stress
1	22	Average to low level of stress
2	17	Average to low level of stress
3	25	Mild to moderate level of stress
4	33	Mild to moderate level of stress
5	14	Average to low level of stress
6	14	Average to low level of stress
7	26	Mild to moderate level of stress
8	21	Average to low level of stress
9	23	Average to low level of stress
10	31	Mild to moderate level of stress
11	17	Average to low level of stress
12	17	Average to low level of stress
13	11	Average to low level of stress
14	18	Average to low level of stress
15	23	Average to low level of stress
16	17	Average to low level of stress
17	16	Average to low level of stress
18	23	Average to low level of stress
19	12	Average to low level of stress
20	16	Average to low level of stress
Mean	19.8	
SD	5.9	
Min	11.0	
Max	33.0	

Table 1.4: Systolic blood pressure of TCOs over an eight hour shift at the workstations selected by Spoornet for the MWL-index validation. Mean and standard deviation of systolic blood pressure values for each TCO was calculated over the whole eight hour shift and from 08:00-14:00. Minimum and maximum values for each TCO are also reported. Mean, standard deviation, minimum and maximum values for every 2 hours from 06:00 – 14:00 are also reported for all the TCOs.

	Systolic blood pressure (mmHg)								
Subject	06:00	08:00	10:00	12:00	14:00	Indv. Mean 06:00-14:00	Indv. Mean 08:00-14:00	Min.	Max.
1	122.7	121.3	115.0	119.7	119.7	119.7 ± 2.9	118.9 ± 2.7	115.0	122.7
2	91.7	89.0	93.7	94.7	95.0	92.8 ± 2.5	93.1 ± 2.8	89.0	95.0
3	129.0	118.7	115.3	122.3	120.7	121.2 ± 5.1	119.3 ± 3.0	115.3	129.0
4	130.3	125.3	122.7	133.7	140.7	130.5 ± 7.1	130.6 ± 8.2	122.7	140.7
5	141.0	130.7	133.7	138.7	148.3	138.5 ± 6.8	137.8 ± 7.7	130.7	148.3
6	114.3	116.0	116.0	117.7	117.3	116.3 ± 1.3	116.8 ± 0.9	114.3	117.7
7	133.7	132.3	136.7	130.7	137.0	134.1 ± 2.7	134.2 ± 3.2	130.7	137.0
8	135.0	121.0	120.3	128.3	128.3	126.6 ± 6.1	124.5 ± 4.4	120.3	135.0
9	121.0	125.7	118.3	114.0	121.3	120.1 ± 4.3	119.8 ± 4.9	114.0	125.7
10	117.3	110.3	111.0	114.3	121.3	114.9 ± 4.6	114.3 ± 5.0	110.3	121.3
11									
12	143.7	144.0	145.0	141.3	145.0	143.8 ± 1.5	143.8 ± 1.7	141.3	145.0
13	125.3	119.7	112.3	119.7	102.7	115.9 ± 8.7	113.6 ± 8.1	102.7	125.3
14	120.3	116.3	118.3	122.0	113.0	118.0 ± 3.5	117.4 ± 3.8	113.0	122.0
15	156.7	137.3	141.0	138.3	158.7	146.4 ± 10.4	143.8 ± 10.0	137.3	158.7
16	137.3	118.3	117.0	108.0	109.0	117.9 ± 11.8	113.1 ± 5.3	108.0	137.3
17	115.0	129.3	125.0	118.3	113.0	120.1 ± 6.9	121.4 ± 7.2	113.0	129.3
18	139.3	150.3	120.7	134.0	132.0	135.3 ± 10.8	134.3 ± 12.2	120.7	150.3
19		109.5	108.3	94.3	111.0	105.8 ± 7.7	105.8 ± 7.7	94.3	111.0
20	108.7	102.0	101.7	102.3	97.0	102.3 ± 4.2	100.8 ± 2.5	97.0	108.7
Mean	126.8	122.0	119.6	120.6	122.7				
SD	15.0	14.3	12.8	14.0	17.5				
Min	91.7	89.0	93.7	94.3	95.0				
Max	156.7	150.3	145.0	141.3	158.7				

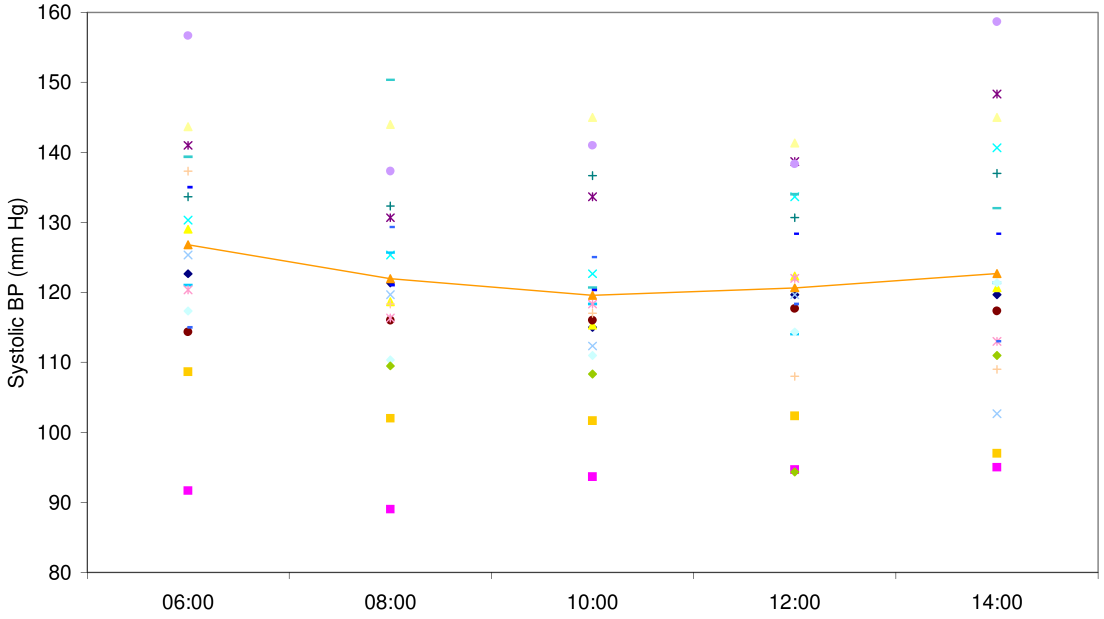


Figure 1.2: Systolic blood pressure for each individual from 06:00 to 14:00.

Table 1.5: Diastolic blood pressure of TCOs over an eight hour shift at the workstations selected by Spoornet for the MWL-index validation. Mean and standard deviation of diastolic blood pressure values for each TCO was calculated over the whole eight hour shift and from 08:00-14:00. Minimum and maximum values for each TCO are also reported. Mean, standard deviation, minimum and maximum values for every 2 hours from 06:00 – 14:00 are also reported for all the TCOs.

	Diastolic blood pressure (mmHg)								
Subject	06:00	08:00	10:00	12:00	14:00	Indv. Mean 06:00-14:00	Indv. Mean 08:00-14:00	Min.	Max.
1	85.7	97.7	77.3	79.7	89.0	85.9 ± 8.1	85.9 ± 9.3	77.3	97.7
2	65.0	65.3	65.0	66.3	69.0	66.1 ± 1.7	66.4 ± 1.8	65.0	69.0
3	105.0	80.3	95.3	87.3	83.7	90.3 ± 9.9	86.7 ± 6.4	80.3	105.0
4	102.0	92.3	92.7	105.0	99.0	98.2 ± 5.6	97.3 ± 6.0	92.3	105.0
5	92.0	99.3	95.7	86.7	103.0	95.3 ± 6.3	96.2 ± 7.0	86.7	103.0
6	82.7	87.7	88.3	75.7	82.3	83.3 ± 5.1	83.5 ± 5.9	75.7	88.3
7	94.0	93.3	96.3	93.3	93.3	94.1 ± 1.3	94.1 ± 1.5	93.3	96.3
8	95.0	89.0	87.3	92.7	90.7	90.9 ± 3.0	89.9 ± 2.3	87.3	95.0
9	87.0	81.7	82.7	79.3	82.3	82.6 ± 2.8	81.5 ± 1.5	79.3	87.0
10	81.3	79.0	78.0	75.3	80.7	78.9 ± 2.4	78.3 ± 2.2	75.3	81.3
11									
12	100.3	97.3	97.7	98.7	99.7	98.7 ± 1.3	98.3 ± 1.1	97.3	100.3
13	87.7	78.7	71.7	77.7	75.0	78.1 ± 6.0	75.8 ± 3.1	71.7	87.7
14	85.0	82.7	84.7	83.7	85.7	84.3 ± 1.2	84.2 ± 1.3	82.7	85.7
15	86.7	78.0	86.7	83.0	81.7	83.2 ± 3.7	82.3 ± 3.6	78.0	86.7
16	92.0	87.0	85.7	75.0	78.3	83.6 ± 6.9	81.5 ± 5.8	75.0	92.0
17	80.0	89.7	87.0	78.3	81.3	83.3 ± 4.8	84.1 ± 5.2	78.3	89.7
18	94.7	106.0	93.0	98.0	87.7	95.9 ± 6.8	96.2 ± 7.8	87.7	106.0
19		75.0	76.7	70.0	74.0	73.9 ± 2.8	73.9 ± 2.8	70.0	76.7
20	66.7	63.3	66.7	62.0	52.3	62.2 ± 5.9	61.1 ± 6.2	52.3	66.7
Mean	87.9	85.4	84.6	82.5	83.6				
SD	10.7	11.2	9.9	11.4	11.7				
Min	65.0	63.3	65.0	62.0	52.3				
Max	105.0	106.0	97.7	105.0	103.0				

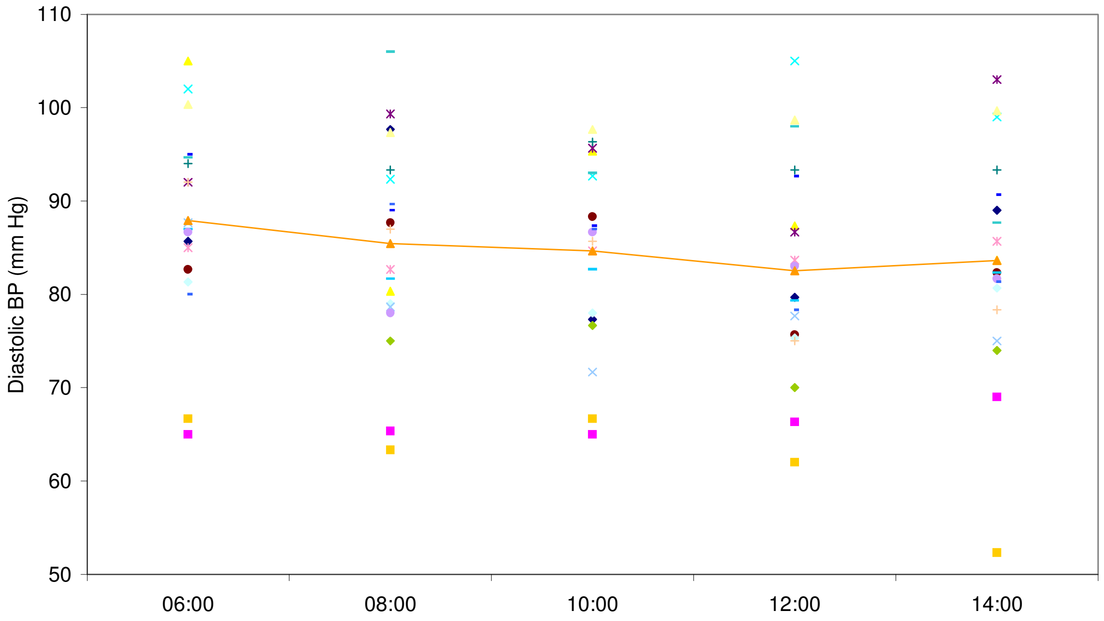


Figure 1.3: Diastolic blood pressure for each individual from 06:00 to 14:00.

Table 1.6: Mean arterial blood pressure of TCOs over an eight hour shift at the workstations selected by Spoornet for the MWL-index validation. Mean and standard deviation of mean arterial blood pressure values for each TCO was calculated over the whole eight hour shift and from 08:00-14:00. Minimum and maximum values for each TCO are also reported. Mean, standard deviation, minimum and maximum values for every 2 hours from 06:00 – 14:00 are also reported for all the TCOs.

	Mean arterial blood pressure (mmHg)								
Subject	06:00	08:00	10:00	12:00	14:00	Indv. Mean 06:00-14:00	Indv. Mean 08:00-14:00	Min.	Max.
1	98.0	105.6	89.9	93.0	99.2	97.1 ± 6.0	96.9 ± 6.9	89.9	105.6
2	73.9	73.2	74.6	75.8	77.7	75.0 ± 1.8	75.3 ± 1.9	73.2	77.7
3	113.0	93.1	102.0	99.0	96.0	100.6 ± 7.7	97.5 ± 3.8	93.1	113.0
4	111.4	103.3	102.7	114.6	112.9	109.0 ± 5.6	108.4 ± 6.2	102.7	114.6
5	108.3	109.8	108.3	104.0	118.1	109.7 ± 5.2	110.1 ± 5.9	104.0	118.1
6	93.2	97.1	97.6	89.7	94.0	94.3 ± 3.2	94.6 ± 3.6	89.7	97.6
7	107.2	106.3	109.8	105.8	107.9	107.4 ± 1.6	107.4 ± 1.8	105.8	109.8
8	108.3	99.7	98.3	104.6	103.2	102.8 ± 4.0	101.4 ± 2.9	98.3	108.3
9	98.3	96.3	94.6	90.9	95.3	95.1 ± 2.7	94.3 ± 2.4	90.9	98.3
10	93.3	89.4	89.0	88.3	94.2	90.9 ± 2.7	90.3 ± 2.7	88.3	94.2
11									
12	114.8	112.9	113.4	112.9	114.8	113.8 ± 1.0	113.5 ± 0.9	112.9	114.8
13	100.2	92.3	85.2	91.7	84.2	90.7 ± 6.4	88.4 ± 4.2	84.2	100.2
14	96.8	93.9	95.9	96.4	94.8	95.6 ± 1.2	95.3 ± 1.1	93.9	96.8
15	110.0	97.8	104.8	101.4	107.3	104.3 ± 4.8	102.8 ± 4.1	97.8	110.0
16	107.1	97.4	96.1	86.0	88.6	95.0 ± 8.3	92.0 ± 5.6	86.0	107.1
17	91.7	102.9	99.7	91.7	91.9	95.6 ± 5.3	96.5 ± 5.6	91.7	102.9
18	109.6	120.8	102.2	110.0	102.4	109.0 ± 7.6	108.9 ± 8.7	102.2	120.8
19		86.5	87.2	78.1	86.3	84.5 ± 4.3	84.5 ± 4.3	78.1	87.2
20	80.7	76.2	78.3	75.4	67.2	75.6 ± 5.1	74.3 ± 4.9	67.2	80.7
Mean	100.9	97.6	96.3	95.2	96.6				
SD	11.3	11.6	10.3	11.8	12.8				
Min	73.9	73.2	74.6	75.4	67.2				
Max	114.8	120.8	113.4	114.6	118.1				

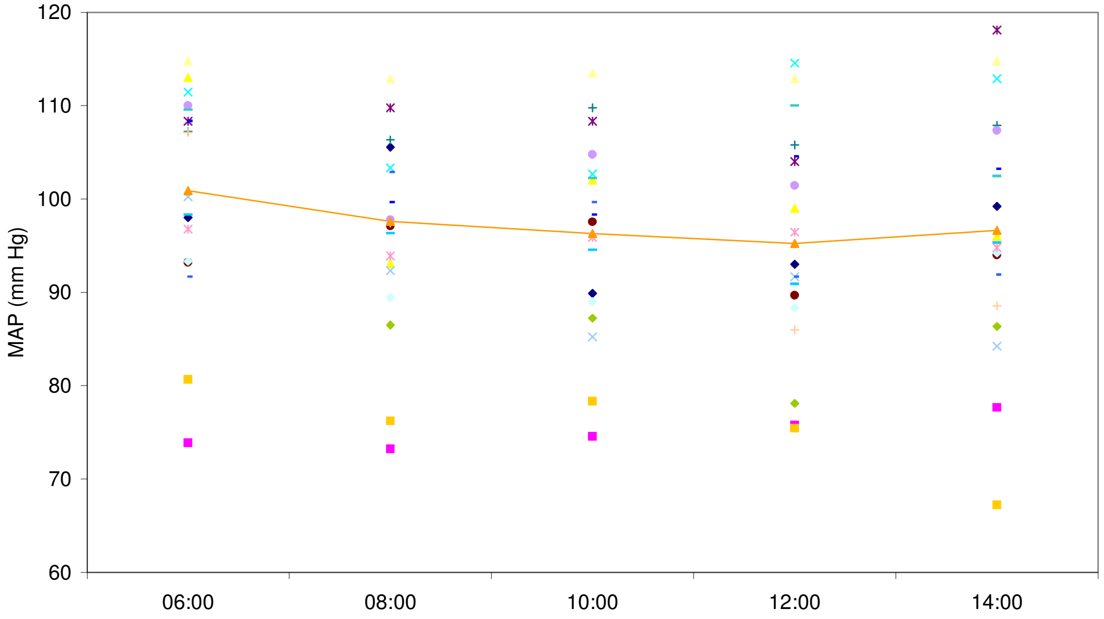


Figure 1.4: Mean arterial blood pressure (MAP) for each individual from 06:00 to 14:00.

Table 1.7: Mean pulse pressure of TCOs over an eight hour shift at the workstations selected by Spoornet for the MWL-index validation. Mean and standard deviation of pulse pressure values for each TCO was calculated over the whole eight hour shift and from 08:00-14:00. Minimum and maximum values for each TCO are also reported. Mean, standard deviation, minimum and maximum values for every 2 hours from 06:00 – 14:00 are also reported for all the TCOs.

	Pulse pressure (mmHg)								
Subject	06:00	08:00	10:00	12:00	14:00	Indv. Mean 06:00-14:00	Indv. Mean 08:00-14:00	Min.	Max.
1	37.0	23.7	37.7	40.0	30.7	33.8 ± 6.6	33.0 ± 7.4	23.7	40.0
2	26.7	23.7	28.7	28.3	26.0	26.7 ± 2.0	26.7 ± 2.3	23.7	28.7
3	24.0	38.3	20.0	35.0	37.0	30.9 ± 8.3	32.6 ± 8.5	20.0	38.3
4	28.3	33.0	30.0	28.7	41.7	32.3 ± 5.5	33.3 ± 5.8	28.3	41.7
5	49.0	31.3	38.0	52.0	45.3	43.1 ± 8.4	41.7 ± 9.0	31.3	52.0
6	31.7	28.3	27.7	42.0	35.0	32.9 ± 5.9	33.3 ± 6.7	27.7	42.0
7	39.7	39.0	40.3	37.3	43.7	40.0 ± 2.3	40.1 ± 2.7	37.3	43.7
8	40.0	32.0	33.0	35.7	37.7	35.7 ± 3.3	34.6 ± 2.6	32.0	40.0
9	34.0	44.0	35.7	34.7	39.0	37.5 ± 4.1	38.3 ± 4.2	34.0	44.0
10	36.0	31.3	33.0	39.0	40.7	36.0 ± 3.9	36.0 ± 4.5	31.3	40.7
11									
12	43.3	46.7	47.3	42.7	45.3	45.1 ± 2.0	45.5 ± 2.1	42.7	47.3
13	37.7	41.0	40.7	42.0	27.7	37.8 ± 5.9	37.8 ± 6.8	27.7	42.0
14	35.3	33.7	33.7	38.3	27.3	33.7 ± 4.0	33.3 ± 4.5	27.3	38.3
15	70.0	59.3	54.3	55.3	77.0	63.2 ± 9.9	61.5 ± 10.6	54.3	77.0
16	45.3	31.3	31.3	33.0	30.7	34.3 ± 6.2	31.6 ± 1.0	30.7	45.3
17	35.0	39.7	38.0	40.0	31.7	36.9 ± 3.5	37.3 ± 3.9	31.7	40.0
18	44.7	44.3	27.7	36.0	44.3	39.4 ± 7.5	38.1 ± 8.0	27.7	44.7
19		34.5	31.7	24.3	37.0	31.9 ± 5.5	31.9 ± 5.5	24.3	37.0
20	42.0	38.7	35.0	40.3	44.7	40.1 ± 3.6	39.7 ± 4.0	35.0	44.7
Mean	38.9	36.5	34.9	38.1	39.1				
SD	10.2	8.5	7.6	7.4	11.2				
Min	24.0	23.7	20.0	24.3	26.0				
Max	70.0	59.3	54.3	55.3	77.0				

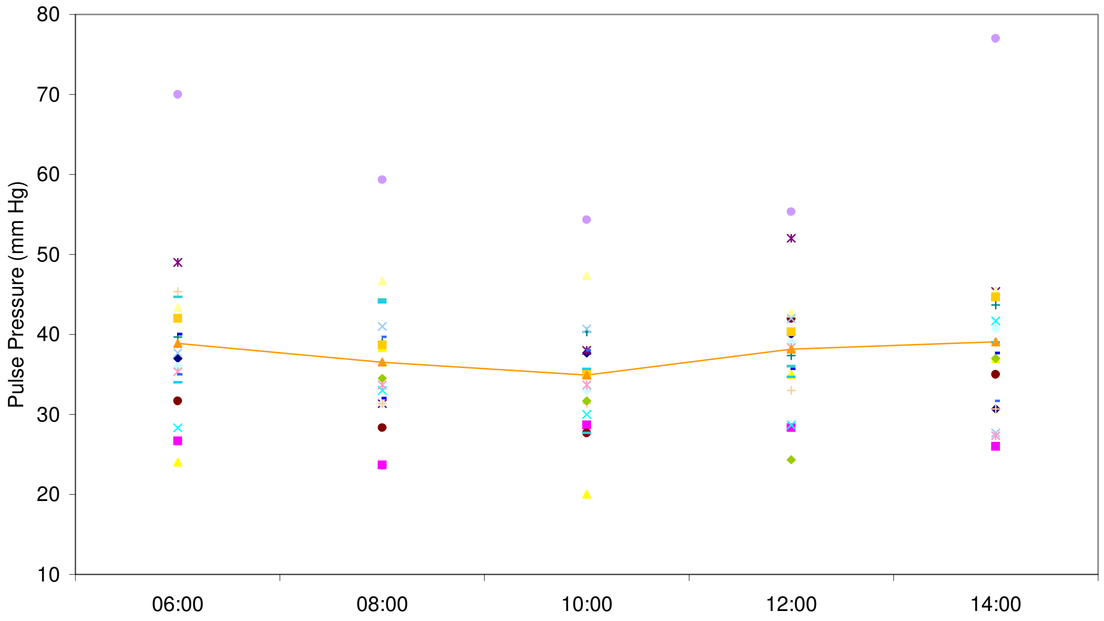


Figure 1.5: Pulse pressure for each individual from 06:00 to 14:00.

Table 1.8: Mean heart rate (measured with the Minimitter) of TCOs over an eight hour shift at the workstations selected by Spoornet for the MWL-index validation. Mean and standard deviation of heart rate values for each TCO was calculated over the whole eight hour shift and from 08:00-14:00. Minimum and maximum values for each TCO are also reported. Mean, standard deviation, minimum and maximum values for every 2 hours from 06:00 – 14:00 are also reported for all the TCOs.

	Heart rate (beats.min^{-1})						
Subject	06:00 - 08:00	08:00 - 10:00	10:00 - 12:00	12:00 – 14:00	Mean	Min.	Max.
1	79.2	76.2	75.0	75.9	76.6 ± 1.8	75.0	79.2
2	89.1	89.3	86.1	85.5	87.5 ± 2.0	85.5	89.3
3	70.2	67.1	64.2	62.9	66.1 ± 3.2	62.9	70.2
4	89.6	88.1	86.5	88.0	88.1 ± 1.3	86.5	89.6
5		81.8	81.5	81.2	81.5 ± 0.3	81.2	81.8
6	82.4	81.9	80.7	83.5	82.1 ± 1.2	80.7	83.5
7	86.5	81.9	82.0	86.5	84.2 ± 2.6	81.9	86.5
8	87.5	69.4	68.5	71.0	74.1 ± 9.0	68.5	87.5
9	87.7	89.5	88.7	88.2	88.5 ± 0.8	87.7	89.5
10	76.6	77.2	78.4	76.2	77.1 ± 1.0	76.2	78.4
11	80.4	72.3	77.9	80.9	77.9 ± 3.9	72.3	80.9
12	89.2	88.5	90.3	90.9	89.7 ± 1.1	88.5	90.9
13	86.3	82.4	79.1	83.4	82.8 ± 3.0	79.1	86.3
14	86.9	85.5	85.5	83.4	85.3 ± 1.4	83.4	86.9
15	67.0	59.7	64.4	64.1	63.8 ± 3.0	59.7	67.0
16	77.5	70.4	68.9	69.4	71.5 ± 4.1	68.9	77.5
17	80.3	81.7	80.9	80.8	80.9 ± 0.6	80.3	81.7
18	95.2	91.2	96.0	97.6	95.0 ± 2.7	91.2	97.6
19	83.3	80.7	81.5	84.3	82.4 ± 1.7	80.7	84.3
20	75.1	71.8	59.7	63.0	67.4 ± 7.2	59.7	75.1
Mean	82.6	79.3	78.8	79.8			
SD	7.2	8.6	9.5	9.6			
Min	67.0	59.7	59.7	62.9			
Max	95.2	91.2	96.0	97.6			

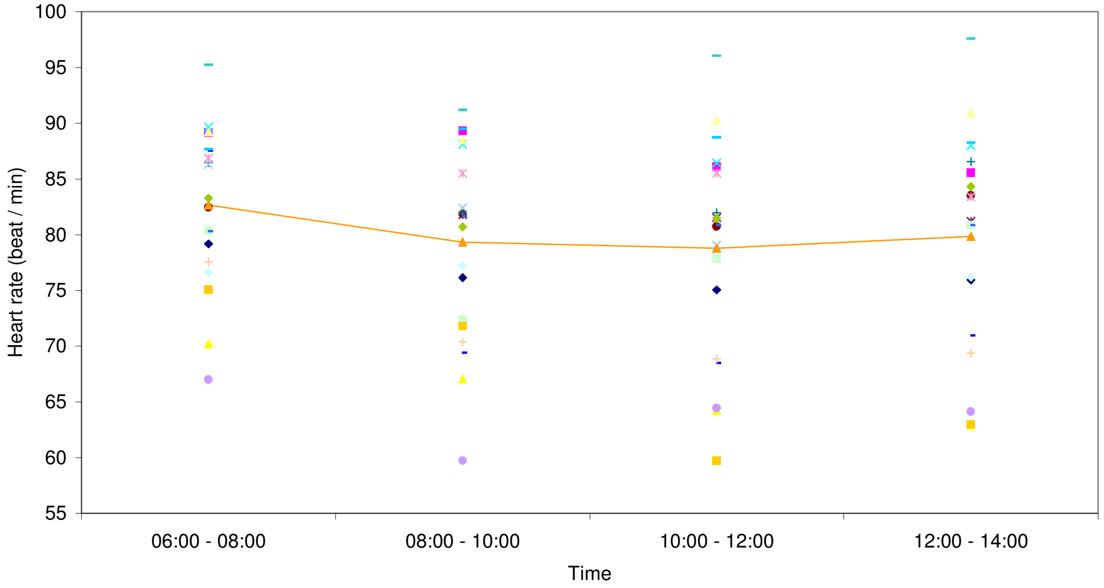


Figure 1.6: Heart rate for each individual from 06:00 to 08:00. (Minimitter)

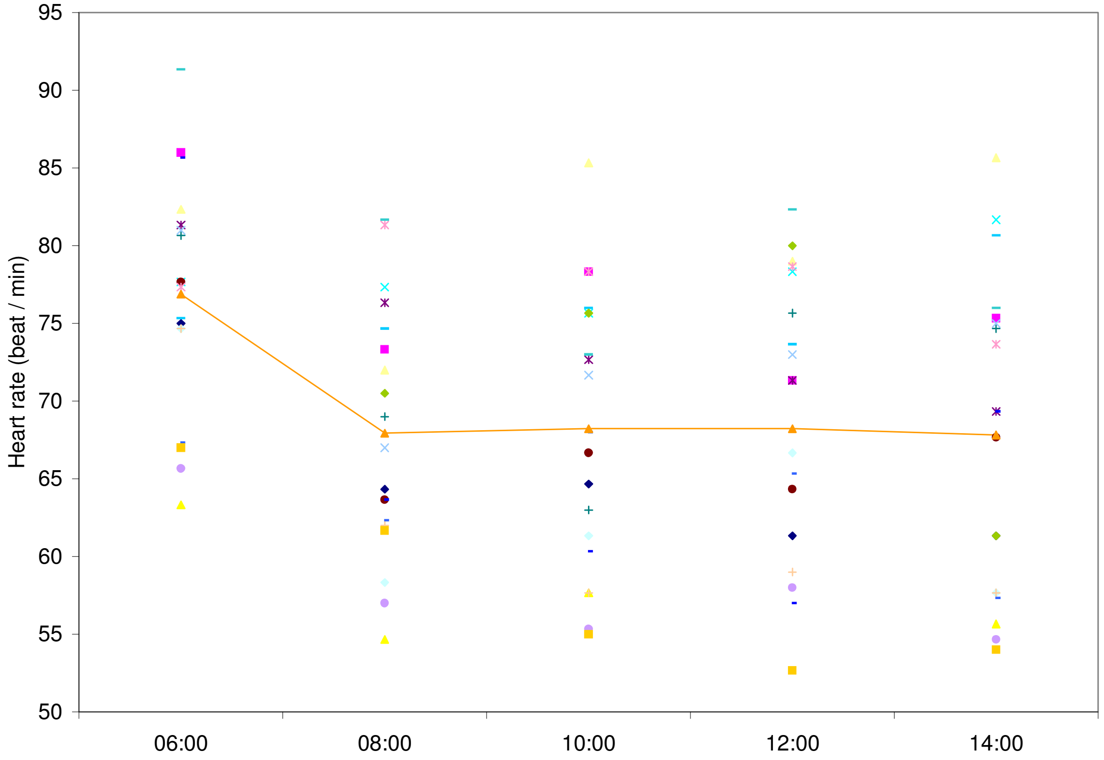


Figure 1.7: Heart rate for each individual from 06:00 to 08:00. (Blood pressure meter).

Table 1.9: Mean RR-intervals of TCOs over an eight hour shift at the workstations selected by Spoornet for the MWL-index validation. Mean and standard deviation of RR-interval values for each TCO was calculated over the whole eight hour shift and from 08:00-14:00. Minimum and maximum values for each TCO are also reported. Mean, standard deviation, minimum and maximum values for every 2 hours from 06:00 – 14:00 are also reported for all the TCOs.

	RR-interval (ms)						
Subject	06:00 - 08:00	08:00 - 10:00	10:00 - 12:00	12:00 - 14:00	Mean	Min.	Max.
1	766.5 ± 52.1	796.3 ± 54.5	808.1 ± 57.7	797.8 ± 55.8	792.2 ± 17.9	766.5	808.1
2	677.3 ± 35.8	677.5 ± 37.7	703.6 ± 40.1	708.9 ± 45.1	691.8 ± 16.8	677.3	708.9
3	875.4 ± 70.4	909.7 ± 82.4	949.2 ± 87.0	972.8 ± 87.3	926.8 ± 43.0	875.4	972.8
4	670.7 ± 16.7	682.8 ± 17.3	696.3 ± 19.1	684.6 ± 22.2	683.6 ± 10.5	670.7	696.3
5		738.6 ± 38.1	741.9 ± 36.9	745.6 ± 40.5	742.0 ± 3.5	738.6	745.6
6	738.5 ± 47.1	744.4 ± 46.6	758.2 ± 51.2	727.3 ± 43.0	742.1 ± 12.9	727.3	758.2
7	701.5 ± 37.2	742.4 ± 47.1	748.6 ± 48.7	699.9 ± 39.0	723.1 ± 26.0	699.9	748.6
8	691.5 ± 31.3	871.1 ± 46.5	882.0 ± 46.4	857.1 ± 44.4	825.4 ± 89.9	691.5	882.0
9	689.5 ± 35.9	675.5 ± 37.0	681.2 ± 37.4	684.3 ± 33.1	682.6 ± 5.8	675.5	689.5
10	791.3 ± 45.1	786.0 ± 50.2	774.0 ± 52.9	797.2 ± 60.3	787.1 ± 9.9	774.0	797.2
11	757.5 ± 32.2	836.9 ± 39.9	789.4 ± 38.6	754.9 ± 32.6	784.7 ± 38.2	754.9	836.9
12	677.4 ± 37.2	683.1 ± 38.0	670.0 ± 41.1	665.5 ± 40.8	674.0 ± 7.8	665.5	683.1
13	698.6 ± 29.1	733.6 ± 38.2	765.9 ± 35.0	723.1 ± 32.6	730.3 ± 27.9	698.6	765.9
14	694.2 ± 36.9	705.6 ± 29.3	705.7 ± 28.3	723.2 ± 29.8	707.2 ± 11.9	694.2	723.2
15	900.0 ± 18.5	1009.3 ± 24.1	935.9 ± 21.1	938.8 ± 20.6	946.0 ± 45.8	900.0	1009.3
16	779.8 ± 30.6	857.4 ± 33.3	879.2 ± 36.6	871.9 ± 37.9	847.1 ± 45.8	779.8	879.2
17	757.5 ± 64.3	745.8 ± 65.4	751.7 ± 62.9	750.9 ± 60.6	751.5 ± 4.8	745.8	757.5
18	633.8 ± 29.2	665.2 ± 30.5	628.9 ± 30.3	618.4 ± 25.1	636.6 ± 20.1	618.4	665.2
19	729.0 ± 49.3	751.5 ± 44.4	744.9 ± 48.4	719.0 ± 46.3	736.1 ± 14.8	719.0	751.5
20	809.6 ± 35.8	844.3 ± 40.6	1015.3 ± 51.8	963.3 ± 48.5	908.1 ± 97.2	809.6	1015.3
Mean	738.9	772.8	781.5	770.2			
SD	70.1	91.1	102.0	101.0			
Min	633.8	665.2	628.9	618.4			
Max	900.0	1009.3	1015.3	972.8			

Table 1.10: Mean low frequency values of TCOs over an eight hour shift at the workstations selected by Spoornet for the MWL-index validation. Mean and standard deviation of low frequency derived from Fast Fourier analysis for each TCO was calculated over the whole eight hour shift and from 08:00-14:00. Minimum and maximum values for each TCO are also reported. Mean, standard deviation, minimum and maximum values for every 2 hours from 06:00 – 14:00 are also reported for all the TCOs.

	LF (ms^2)						
Subject	06:00 - 08:00	08:00 - 10:00	10:00 - 12:00	12:00 - 14:00	Mean	Min.	Max.
1	1173.8	1163.9	1380.2	1331.7	1262.4 ± 109.9	1163.9	1380.2
2	332.0	421.1	426.8	540.0	430.0 ± 85.2	332.0	540.0
3	1679.2	2377.9	2573.9	2552.9	2296.0 ± 420.4	1679.2	2573.9
4	107.6	112.0	136.6	207.2	140.9 ± 46.0	107.6	207.2
5		563.5	559.9	669.1	597.5 ± 62.0	559.9	669.1
6	780.3	723.4	820.3	587.6	727.9 ± 101.6	587.6	820.3
7	427.4	657.1	737.6	513.8	583.9 ± 139.5	427.4	737.6
8	408.8	727.0	767.5	690.0	648.3 ± 162.8	408.8	767.5
9	434.6	519.7	564.6	397.7	479.1 ± 76.5	397.7	564.6
10	773.8	895.9	1054.3	1397.8	1030.4 ± 270.5	773.8	1397.8
11	327.1	524.2	573.4	401.7	456.6 ± 112.5	327.1	573.4
12	508.8	503.5	657.9	669.4	584.9 ± 91.1	503.5	669.4
13	242.3	403.5	338.5	264.9	312.3 ± 73.4	242.3	403.5
14	645.6	364.6	327.4	378.7	429.1 ± 146.0	327.4	645.6
15	75.1	111.0	88.7	98.1	93.2 ± 15.2	75.1	111.0
16	310.7	357.4	434.4	453.2	388.9 ± 66.6	310.7	453.2
17	1261.5	1238.9	1043.5	1127.0	1167.7 ± 101.6	1043.5	1261.5
18	362.2	382.8	414.8	252.9	353.2 ± 70.3	252.9	414.8
19	835.0	698.3	792.7	733.2	764.8 ± 60.9	698.3	835.0
20	371.6	476.7	669.3	571.5	522.3 ± 127.6	371.6	669.3
Mean	582.0	661.1	718.1	691.9			
SD	417.1	496.9	537.2	558.7			
Min	75.1	111.0	88.7	98.1			
Max	1679.2	2377.9	2573.9	2552.9			

Table 1.11, the LF to HF ratios in Table 1.12, the RMSSD value in Table 1.13 and the graphical presentations for each individual, over 15 minute intervals, over the total work session in Figures1.8 to 1.12.

Table 1.11: Mean high frequency values of TCOs over an eight hour shift at the workstations selected by Spoornet for the MWL-index validation. Mean and standard deviation of high frequency derived from Fast Fourier analysis for each TCO was calculated over the whole eight hour shift and from 08:00-14:00. Minimum and maximum values for each TCO are also reported. Mean, standard deviation, minimum and maximum values for every 2 hours from 06:00 – 14:00 are also reported for all the TCOs.

	HF (ms^2)						
Subject	06:00 - 08:00	08:00 - 10:00	10:00 - 12:00	12:00 - 14:00	Mean	Min.	Max.
1	93.6	128.6	130.8	118.7	117.9 ± 17.0	93.6	130.8
2	143.9	183.6	252.9	296.5	219.2 ± 68.4	143.9	296.5
3	592.8	786.0	746.4	918.2	760.9 ± 134.0	592.8	918.2
4	16.4	19.7	25.1	36.1	24.3 ± 8.6	16.4	36.1
5		100.3	69.5	76.1	82.0 ± 16.2	69.5	100.3
6	201.3	209.9	249.5	211.9	218.2 ± 21.4	201.3	249.5
7	154.6	309.9	320.7	140.7	231.5 ± 97.1	140.7	320.7
8	85.4	221.3	184.4	189.0	170.0 ± 58.8	85.4	221.3
9	176.0	159.4	115.4	127.8	144.6 ± 27.9	115.4	176.0
10	200.7	286.4	341.3	383.0	302.8 ± 78.8	200.7	383.0
11	91.7	165.3	102.5	53.8	103.3 ± 46.3	53.8	165.3
12	129.5	134.0	114.4	111.9	122.5 ± 11.0	111.9	134.0
13	111.8	237.2	171.9	162.9	170.9 ± 51.5	111.8	237.2
14	26.3	21.7	25.9	26.6	25.1 ± 2.3	21.7	26.6
15	49.5	98.6	58.8	51.5	64.6 ± 23.0	49.5	98.6
16	58.9	80.9	96.5	125.8	90.5 ± 28.1	58.9	125.8
17	693.3	711.9	746.2	735.1	721.6 ± 23.7	693.3	746.2
18	32.0	45.0	38.3	26.1	35.3 ± 8.1	26.1	45.0
19	356.3	237.4	310.3	227.5	282.9 ± 61.3	227.5	356.3
20	146.9	215.3	445.5	387.7	298.8 ± 140.8	146.9	445.5
Mean	176.9	217.6	227.3	220.4			
SD	183.2	199.7	212.6	235.0			
Min	16.4	19.7	25.1	26.1			
Max	693.3	786.0	746.4	918.2			

Table 1.12: Mean Total (low plus high) frequency values of TCOs over an eight hour shift at the workstations selected by Spoornet for the MWL-index validation. Mean and standard deviation of total frequency derived from Fast Fourier analysis for each TCO was calculated over the whole eight hour shift and from 08:00-14:00. Minimum and maximum values for each TCO are also reported. Mean, standard deviation, minimum and maximum values for every 2 hours from 06:00 – 14:00 are also reported for all the TCOs.

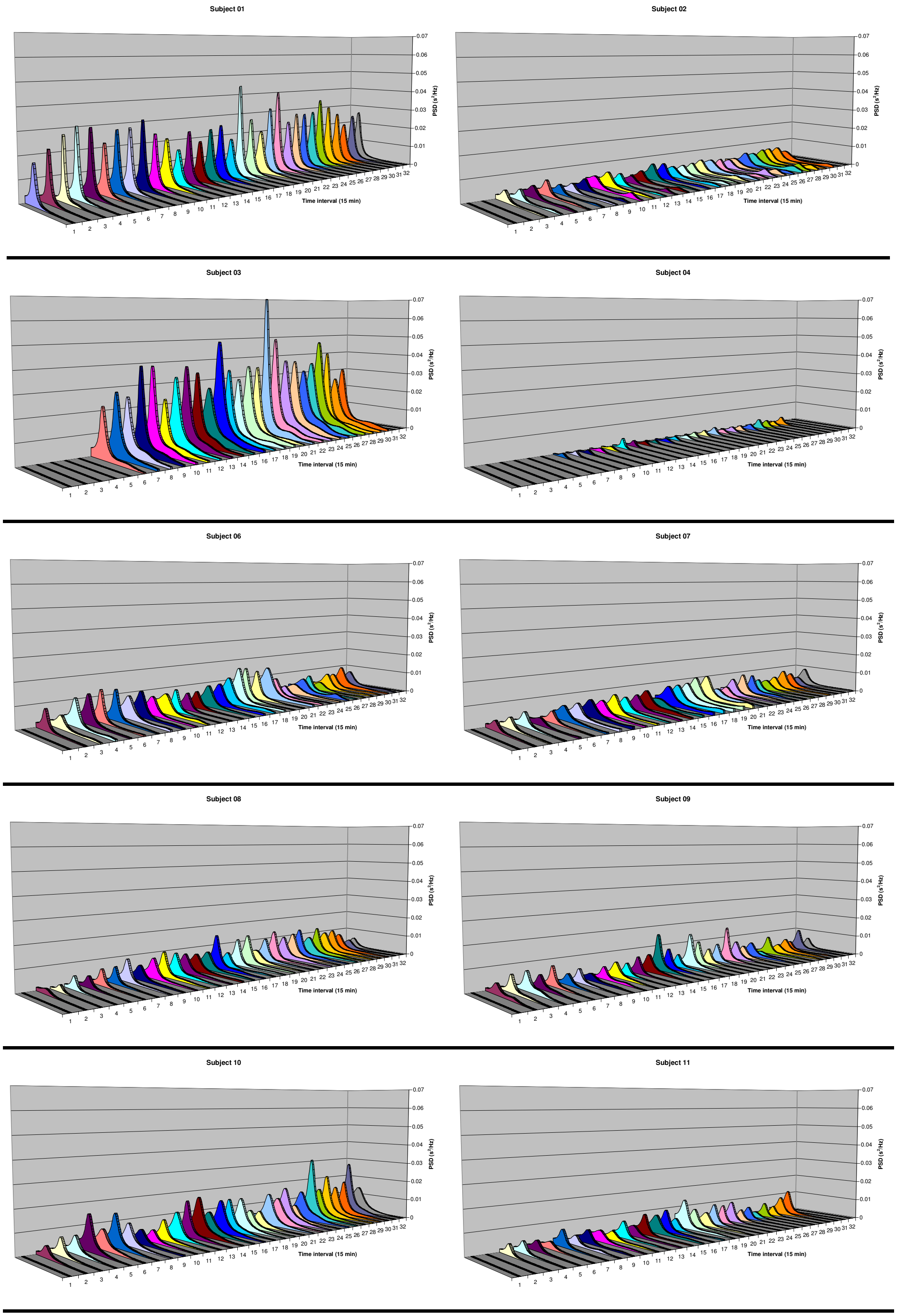
	Total (ms^2)						
Subject	06:00 - 08:00	08:00 - 10:00	10:00 - 12:00	12:00 - 14:00	Mean	Min.	Max.
1	1267.5	1292.5	1510.9	1450.5	1380.3 ± 118.9	1267.5	1510.9
2	475.9	604.7	679.6	836.5	649.2 ± 150.6	475.9	836.5
3	2272.0	3163.9	3320.2	3471.1	3056.8 ± 538.0	2272.0	3471.1
4	124.0	131.7	161.7	243.3	165.2 ± 54.5	124.0	243.3
5		663.8	629.4	745.2	509.6 ± 343.2	0.0	745.2
6	981.6	933.3	1069.8	799.6	946.1 ± 112.8	799.6	1069.8
7	581.9	967.0	1058.2	654.4	815.4 ± 232.6	581.9	1058.2
8	494.2	948.2	951.9	879.0	818.4 ± 218.7	494.2	951.9
9	610.6	679.1	680.0	525.5	623.8 ± 73.2	525.5	680.0
10	974.5	1182.3	1395.6	1780.8	1333.3 ± 344.3	974.5	1780.8
11	418.9	689.4	675.9	455.5	559.9 ± 142.6	418.9	689.4
12	638.4	637.5	772.3	781.3	707.4 ± 80.3	637.5	781.3
13	354.1	640.7	510.4	427.8	483.3 ± 122.9	354.1	640.7
14	671.9	386.3	353.3	405.3	454.2 ± 146.7	353.3	671.9
15	124.6	209.6	147.4	149.6	157.8 ± 36.4	124.6	209.6
16	369.6	438.3	530.9	579.0	479.5 ± 93.6	369.6	579.0
17	1954.7	1950.8	1789.7	1862.1	1889.3 ± 79.0	1789.7	1954.7
18	394.2	427.7	453.1	279.0	388.5 ± 76.9	279.0	453.1
19	1191.3	935.7	1103.0	960.7	1047.7 ± 120.8	935.7	1191.3
20	518.5	692.0	1114.8	959.2	821.1 ± 266.8	518.5	1114.8
Mean	720.9	878.7	945.4	912.3			
SD	583.1	674.6	707.3	761.6			
Min	0.0	131.7	147.4	149.6			
Max	2272.0	3163.9	3320.2	3471.1			

Table 1.13: Mean Ratio (LF/HF) values of TCOs over an eight hour shift at the workstations selected by Spoornet for the MWL-index validation. Mean and standard deviation of Ratio (LF/HF) derived from Fast Fourier analysis for each TCO was calculated over the whole eight hour shift and from 08:00-14:00. Minimum and maximum values for each TCO are also reported. Mean, standard deviation, minimum and maximum values for every 2 hours from 06:00 – 14:00 are also reported for all the TCOs.

	Ratio (LF/HF)						
Subject	06:00 - 08:00	08:00 - 10:00	10:00 - 12:00	12:00 – 14:00	Mean	Min.	Max.
1	12.5	9.1	10.6	11.2	10.8 ± 1.4	9.1	12.5
2	2.3	2.3	1.7	1.8	2.0 ± 0.3	1.7	2.3
3	2.8	3.0	3.4	2.8	3.0 ± 0.3	2.8	3.4
4	6.6	5.7	5.4	5.7	5.9 ± 0.5	5.4	6.6
5		5.6	8.1	8.8	7.5 ± 1.7	5.6	8.8
6	3.9	3.4	3.3	2.8	3.3 ± 0.5	2.8	3.9
7	2.8	2.1	2.3	3.7	2.7 ± 0.7	2.1	3.7
8	4.8	3.3	4.2	3.7	4.0 ± 0.7	3.3	4.8
9	2.5	3.3	4.9	3.1	3.4 ± 1.0	2.5	4.9
10	3.9	3.1	3.1	3.6	3.4 ± 0.4	3.1	3.9
11	3.6	3.2	5.6	7.5	5.0 ± 2.0	3.2	7.5
12	3.9	3.8	5.8	6.0	4.9 ± 1.2	3.8	6.0
13	2.2	1.7	2.0	1.6	1.9 ± 0.2	1.6	2.2
14	24.5	16.8	12.6	14.2	17.1 ± 5.3	12.6	24.5
15	1.5	1.1	1.5	1.9	1.5 ± 0.3	1.1	1.9
16	5.3	4.4	4.5	3.6	4.4 ± 0.7	3.6	5.3
17	1.8	1.7	1.4	1.5	1.6 ± 0.2	1.4	1.8
18	11.3	8.5	10.8	9.7	10.1 ± 1.3	8.5	11.3
19	2.3	2.9	2.6	3.2	2.8 ± 0.4	2.3	3.2
20	2.5	2.2	1.5	1.5	1.9 ± 0.5	1.5	2.5
Mean	5.3	4.4	4.8	4.9			
SD	5.5	3.6	3.4	3.6			
Min	1.5	1.1	1.4	1.5			
Max	24.5	16.8	12.6	14.2			

Table 1.14: Mean RMSSD values of TCOs over an eight hour shift at the workstations selected by Spoornet for the MWL-index validation. Mean and standard deviation of RMSSD derived from Fast Fourier analysis for each TCO was calculated over the whole eight hour shift and from 08:00-14:00. Minimum and maximum values for each TCO are also reported. Mean, standard deviation, minimum and maximum values for every 2 hours from 06:00 – 14:00 are also reported for all the TCOs.

	RMSSD						
Subject	06:00 - 08:00	08:00 - 10:00	10:00 - 12:00	12:00 - 14:00	Mean	Min.	Max.
1	25.9	29.8	31.9	30.1	29.4 ± 2.5	25.9	31.9
2	32.2	30.1	36.6	41.2	35.0 ± 4.9	30.1	41.2
3	60.2	70.7	72.5	79.4	70.7 ± 7.9	60.2	79.4
4	10.0	10.2	11.9	13.5	11.4 ± 1.6	10.0	13.5
5		22.8	20.6	22.7	22.0 ± 1.3	20.6	22.8
6	34.6	33.6	38.4	33.3	35.0 ± 2.3	33.3	38.4
7	31.6	43.0	44.0	31.2	37.5 ± 7.0	31.2	44.0
8	20.0	37.4	35.5	34.5	31.8 ± 8.0	20.0	37.4
9	25.3	24.0	22.3	22.6	23.6 ± 1.4	22.3	25.3
10	32.0	38.5	39.8	43.6	38.5 ± 4.8	32.0	43.6
11	24.4	32.1	26.2	20.7	25.8 ± 4.8	20.7	32.1
12	23.6	24.5	23.6	23.1	23.7 ± 0.6	23.1	24.5
13	30.3	40.8	36.7	34.2	35.5 ± 4.4	30.3	40.8
14	14.8	12.8	13.6	13.7	13.7 ± 0.8	12.8	14.8
15	20.5	27.7	23.5	21.8	23.4 ± 3.1	20.5	27.7
16	19.4	23.7	25.9	28.6	24.4 ± 3.9	19.4	28.6
17	60.0	60.3	58.3	53.0	57.9 ± 3.4	53.0	60.3
18	12.1	15.5	12.6	10.8	12.7 ± 2.0	10.8	15.5
19	42.5	35.4	40.1	34.6	38.1 ± 3.8	34.6	42.5
20	28.9	34.7	52.4	48.9	41.2 ± 11.2	28.9	52.4
Mean	28.9	32.4	33.3	32.1			
SD	13.6	14.5	15.6	16.0			
Min	10.0	10.2	11.9	10.8			
Max	60.2	70.7	72.5	79.4			


Subject 01

Subject 02

Subject 03

Subject 04

Subject 06

Subject 07

Subject 08

Subject 09

Subject 10

Subject 11

Figure 1.8: Autoregressive analysis of heart rate variability at 15 minute intervals for S1 – S11 from 06:00 to 14:00.

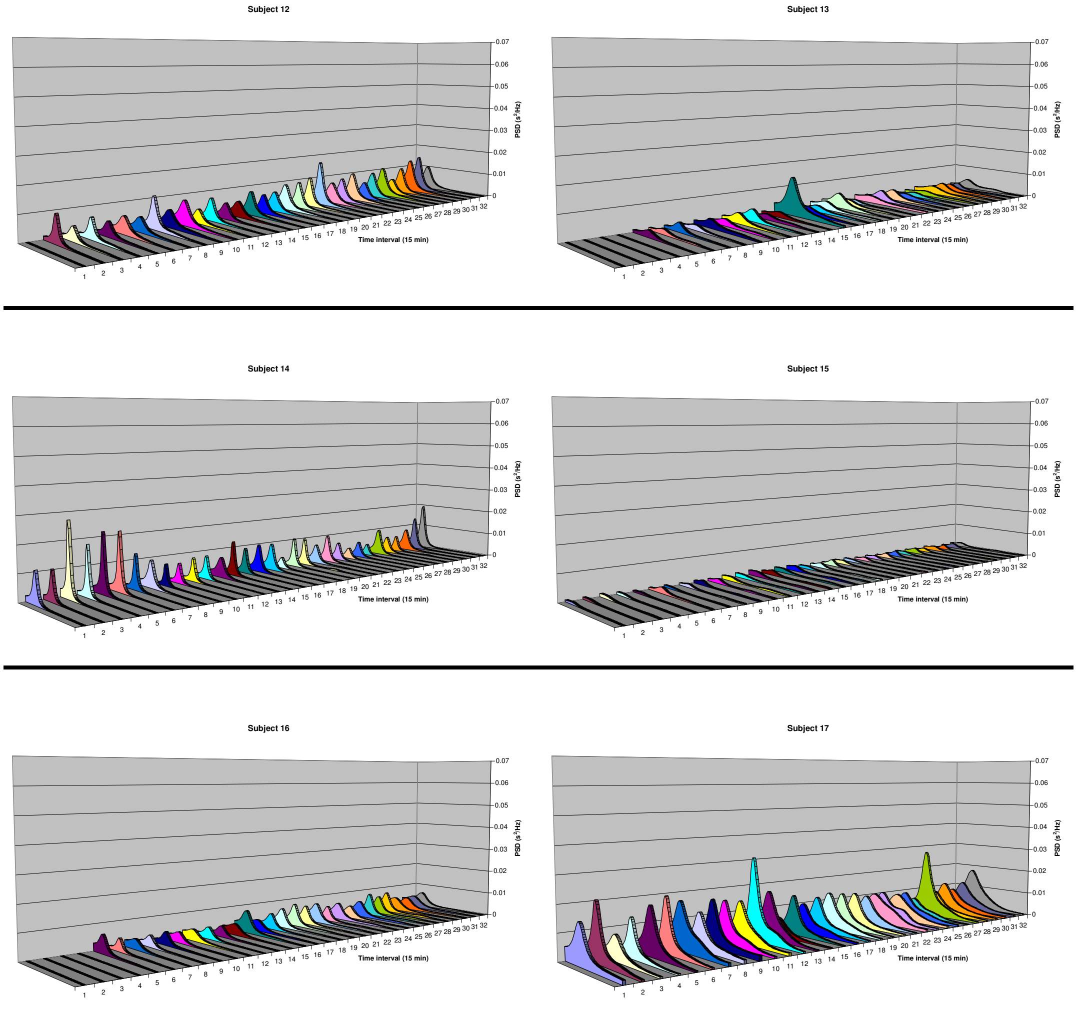


Figure 1.9: Autoregressive analysis of heart rate variability at 15 minute intervals for S12 – S17 from 06:00 to 14:00.

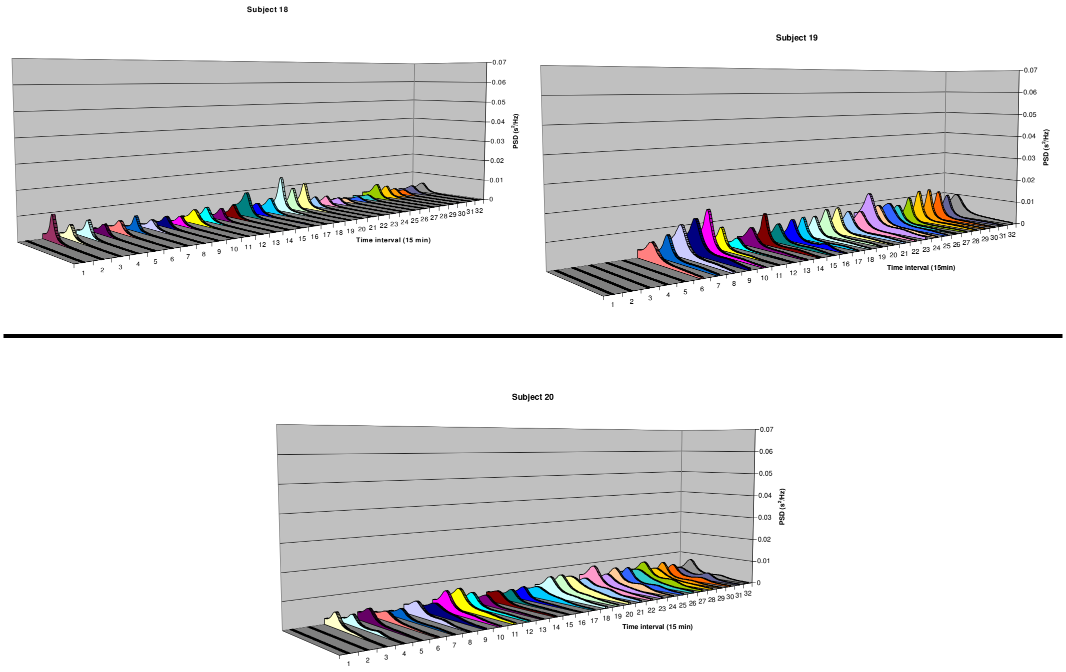


Figure 1.10: Autoregressive analysis of heart rate variability at 15 minute intervals for S18 – S20 from 06:00 to 14:00.

Online time analyses over the work shift can be seen from Tables 1.15 to Table 1.19 with Table 1.15 giving the number of radio communications, Table 1.16 the number of telephonic communications, Table 1.17 the number of times updating the train diagram (schedules), Table 1.18 the number of trains and Figure 1.19 the number of authorizations.

Table 1.15: Number of radio communications at the workstations selected by Spoornet for the MWL-index validation. Total and standard deviation of the number of radio communications over the whole eight hour shift and from 08:00-14:00 and minimum and maximum values for each TCO are also reported. Mean, standard deviation, minimum and maximum values for every 2 hours from 06:00 – 14:00 are also reported for all the TCOs.

	Number of radio communications							
Subject	06:00 - 08:00	08:00 - 10:00	10:00 - 12:00	12:00 – 14:00	Total	Stdev	Min.	Max.
1	21	37	24	32	114	7.3	21	37
2	6	17	14	4	41	6.2	4	17
3	11	17	11	5	44	4.9	5	17
4	5	15	20	18	58	6.7	5	20
5	30	46	62	92	230	26.5	30	92
6	2	10	8	23	43	8.8	2	23
7	21	22	2	35	80	13.6	2	35
8	17	13	25	38	93	11.0	13	38
9	55	18	0	0	73	25.9	0	55
10	47	47	39	37	170	5.3	37	47
11	54	42	27	31	154	12.1	27	54
12	30	40	39	12	121	13.0	12	40
13	3	6	3	3	15	1.5	3	6
14	36	35	36	57	164	10.7	35	57
15	19	19	28	19	85	4.5	19	28
16	5	9	16	11	41	4.6	5	16
17	32	56	49	48	185	10.1	32	56
18	93	68	75	61	297	13.7	61	93
19	40	53	10	48	151	19.3	10	53
20	34	26	34	30	124	3.8	26	34
Mean	28.1	29.8	26.1	30.2				
SD	22.6	18.0	20.0	23.2				
Min	2	6	0	0				
Max	93	68	75	92				

Table 1.16: Number of telephonic communications at the workstations selected by Spoornet for the MWL-index validation. Total and standard deviation of the number of telephonic communications over the whole eight hour shift and from 08:00-14:00 and minimum and maximum values for each TCO are also reported. Mean, standard deviation, minimum and maximum values for every 2 hours from 06:00 – 14:00 are also reported for all the TCOs.

	Number of telephonic communications							
Subject	06:00 - 08:00	08:00 - 10:00	10:00 - 12:00	12:00 - 14:00	Total	Stdev	Min.	Max.
1	4	8	11	13	36	3.9	4	13
2	2	6	2	3	13	1.9	2	6
3	4	21	14	14	53	7.0	4	21
4	10	10	11	6	37	2.2	6	11
5	22	19	18	18	77	1.9	18	22
6	4	7	1	8	20	3.2	1	8
7	2	2	8	7	19	3.2	2	8
8	8	9	4	6	27	2.2	4	9
9	11	7	1	2	21	4.6	1	11
10	13	11	13	10	47	1.5	10	13
11	23	20	24	14	81	4.5	14	24
12	18	11	7	2	38	6.8	2	18
13	8	9	8	10	35	1.0	8	10
14	5	6	9	6	26	1.7	5	9
15	9	10	8	11	38	1.3	8	11
16	0	4	2	1	7	1.7	0	4
17	17	16	11	23	67	4.9	11	23
18	9	8	14	7	38	3.1	7	14
19	6	10	14	6	36	3.8	6	14
20	11	12	8	12	43	1.9	8	12
Mean	9.3	10.3	9.4	9.0				
SD	6.5	5.1	5.9	5.6				
Min	0	2	1	1				
Max	23	21	24	23				

Table 1.17: Number of schedules (updating the train diagram) at the workstations selected by Spoornet for the MWL-index validation. Total and standard deviation of the number of schedules over the whole eight hour shift and from 08:00-14:00 and minimum and maximum values for each TCO are also reported. Mean, standard deviation, minimum and maximum values for every 2 hours from 06:00 – 14:00 are also reported for all the TCOs.

	Number of schedules							
Subject	06:00 - 08:00	08:00 - 10:00	10:00 - 12:00	12:00 - 14:00	Total	Stdev	Min.	Max.
1	16	18	15	16	65	1.3	15	18
2	3	6	8	2	19	2.8	2	8
3	4	9	2	2	17	3.3	2	9
4	4	6	5	6	21	1.0	4	6
5	10	13	17	19	59	4.0	10	19
6	1	2	3	4	10	1.3	1	4
7	7	9	1	13	30	5.0	1	13
8	7	7	10	15	39	3.8	7	15
9	25	14	0	0	39	12.1	0	25
10	17	18	17	21	73	1.9	17	21
11	20	20	15	14	69	3.2	14	20
12	22	25	25	11	83	6.7	11	25
13	2	5	2	2	11	1.5	2	5
14	23	21	25	29	98	3.4	21	29
15	8	5	11	7	31	2.5	5	11
16	3	4	8	5	20	2.2	3	8
17	12	28	27	27	94	7.7	12	28
18	37	27	33	16	113	9.1	16	37
19	14	17	4	13	48	5.6	4	17
20	14	15	14	13	56	0.8	13	15
Mean	12.5	13.5	12.1	11.8				
SD	9.5	8.1	9.7	8.3				
Min	1	2	0	0				
Max	37	28	33	29				

Table 1.18: Number of trains at the workstations selected by Spoornet for the MWL-index validation. Total and standard deviation of the number of trains over the whole eight hour shift and from 08:00-14:00 and minimum and maximum values for each TCO are also reported. Mean, standard deviation, minimum and maximum values for every 2 hours from 06:00 – 14:00 are also reported for all the TCOs.

	Number of trains							
Subject	06:00 - 08:00	08:00 - 10:00	10:00 - 12:00	12:00 - 14:00	Total	Stdev	Min.	Max.
1	4	5	5	4	18	0.6	4	5
2	1	2	2	1	6	0.6	1	2
3	1	2	0	2	5	1.0	0	2
4	0	2	3	1	6	1.3	0	3
5	4	3	6	5	18	1.3	3	6
6	1	0	1	1	3	0.5	0	1
7	4	0	0	6	10	3.0	0	6
8	2	2	2	6	12	2.0	2	6
9	7	4	0	0	11	3.4	0	7
10	7	7	6	6	26	0.6	6	7
11	8	7	5	3	23	2.2	3	8
12	17	16	16	12	61	2.2	12	17
13	2	2	1	2	7	0.5	1	2
14	4	4	5	7	20	1.4	4	7
15	3	4	6	4	17	1.3	3	6
16	1	2	2	1	6	0.6	1	2
17	8	11	10	8	37	1.5	8	11
18	21	22	20	20	83	1.0	20	22
19	11	3	2	4	20	4.1	2	11
20	6	5	6	5	22	0.6	5	6
Mean	5.6	5.2	4.9	4.9				
SD	5.5	5.5	5.2	4.6				
Min	0	0	0	0				
Max	21	22	20	20				

Table 1.19: Number of authorisations at the workstations selected by Spoornet for the MWL-index validation. Total and standard deviation of the number of authorisations over the whole eight hour shift and from 08:00-14:00 and minimum and maximum values for each TCO are also reported. Mean, standard deviation, minimum and maximum values for every 2 hours from 06:00 – 14:00 are also reported for all the TCOs.

	Number of authorisations							
Subject	06:00 - 08:00	08:00 - 10:00	10:00 - 12:00	12:00 - 14:00	Total	Stdev	Min.	Max.
1	5	8	4	5	22	1.7	4	8
2	1	2	2	1	6	0.6	1	2
3	1	2	0	2	5	1.0	0	2
4	0	2	3	1	6	1.3	0	3
5	4	3	6	9	22	2.6	3	9
6	1	2	1	1	5	0.5	1	2
7	4	0	0	6	10	3.0	0	6
8	4	2	2	9	17	3.3	2	9
9	12	4	0	0	16	5.7	0	12
10	8	7	9	7	31	1.0	7	9
11	11	8	7	3	29	3.3	3	11
12	19	11	8	4	42	6.4	4	19
13	2	2	1	2	7	0.5	1	2
14	7	4	6	12	29	3.4	4	12
15	2	1	1	2	6	0.6	1	2
16	1	1	1	1	4	0.0	1	1
17	3	7	7	8	25	2.2	3	8
18	2	2	0	3	7	1.3	0	3
19	20	8	3	7	38	7.3	3	20
20	6	5	6	5	22	0.6	5	6
Mean	5.7	4.1	3.4	4.4				
SD	5.8	3.1	3.0	3.4				
Min	0	0	0	0				
Max	20	11	9	12				

3.2 Comparisons Between Spoornet Models of Workload at the Different Venues and the Physiological Indicators of Individuals Working at the Respective Venues (Validation of the MWL-index)

In this section the MWL-index as received from Spoornet was tested to see whether it reflects the load as calculated by the physiological stress indicators of the individuals. In other words, it investigated whether a correlation exists between division of the different venues into high and low work stress venues by using the MWL-index on the one hand, and the reflection of the workload as reflected by values of the physiological stress indicators tested in the study on the other.

The developed MWL-index as received from Spoornet (using data collected in 2003), hence referred to as Model 1, is seen in Table 2.1, showing the number of SIMS/ETD, the number of authorisations, weighted number of authorisations, number of telephone and radio communications, number of weighted communications, total number of actions, total number of actions, moderators, the final workload index of this model and the original Spoornet rating according to this model and the workstation system.

Table 2.1: Spoornet MWL-index (using data collected in 2003, i.e., Model 1)

Id	SIMS	ETA/ETD	No of data transactions (SIMS + ETA/ETD)	No of auth	Weighted no of auth	No of other tel/radio comms	Weighted no of comms	Total no of actions (weighted)	Mental work-load Index (Model 1)	Spoornet rating (Model 1)	System
1	5	0	5	34	510	204	1020	1535	2629	High	TWS 2
2	0	0	0	9	135	50	250	385	553	Low	TWS 1
3	3	13	16	15	225	100	500	741	1155	Low	RTO 1
4	0	0	0	9	135	9	45	180	281	Low	RTO 2
5	4	0	4	38	570	271	1355	1929	3320	High	TWS 2
6	3	0	3	20	300	130	650	953	1476	Medium	TWS 1
7	0	0	0	35	525	146	730	1255	2228	Low	RTO 1
8	4	12	16	97	1455	345	1725	3196	5789	High	RTO 1
9	4	20	24	39	585	297	1485	2094	3722	High	TWS 1
10	6	12	18	19	285	366	1830	2133	3666	Medium	TWS 3
11	7	0	7	45	675	325	1625	2307	5174	High	RTO 1
12	5	0	5	13	195	264	1320	1520	2214	Low	RTO 1
13	0	0	0	6	90	8	40	130	171	Low	RTO 1
14	0	0	0	42	630	327	1635	2265	4399	High	TWS 1
15	0	10	10	8	120	23	115	245	374	Low	RTO 4
16	0	20	20	31	465	261	1305	1790	3086	Low	RTO 1
17	5	4	9	8	120	88	440	569	866	Low	RTO 1
18	0	0	0	8	120	91	455	575	863	Low	RTO 1
19	2	0	2	46	690	222	1110	1802	3443	High	TWS 1
20	0	0	0	71	1065	82	410	1475	2828	High	TWS 2

SIMS = Spoornet information management system; ETA = Estimated time of arrival; ETD = Estimated time of departure

The above model (Model 1) was of a historical nature and represented workload information not necessary appropriate for the individual presently involved in working at each specific workstation. A revised Spoornet MWL-index (Model 2) was calculated based on real time activities (activities recorded at the time of the physiological measurements). This revised MWL-index (Model 2) is seen in Table 2.2. Table 2.2 shows the calculated MWL-index for Model 1 and Model 2.

Table 2.2: Model 1 (historical data) and Model 2 (real time data recorded at time of physiological measurements). Comparison between Model 1 and Model 2 values indicate no significant differences between the two groups (p = 0.6753). A significant correlation exists between Model 1 and Model 2 values (r = 0.5224; p = 0.0181)

Id	SIMS	ETA/ETD	No of data transactions (SIMS + ETA/ETD)	No of auth	Weighted no of auth	No of other tel/radio comms	Weighted no of comms	Total no of actions (weighted)	Mental work-load Index (Model 2)	50 – 50 Split (Model 2)	Mental work-load Index (Model 1)	Spoornet rating (Model 1)	System
1	0	0	0	22	330	150	750	1080	1797	Low	2629	Medium	TWS 1
2	0	0	0	6	90	56	279	369	561	Low	553	Low	TWS 1
3	1	0	1	6	86	115	575	662	1032	Low	1155	Low	RTO 1
4	0	0	0	6	90	109	543	633	902	Low	281	Low	RTO 2
5	1	0	1	22	330	307	1535	1866	3789	High	3320	High	TWS 2
6	0	0	0	5	75	63	315	390	575	Low	1476	Medium	TWS 1
7	1	0	1	10	150	99	495	646	1147	Low	2228	Low	RTO 1
8	0	0	0	17	255	120	600	855	1724	Low	5789	High	RTO 1
9	0	4	4	32	480	188	940	1424	2458	High	3722	High	TWS 1
10	1	8	9	31	465	217	1085	1559	2601	High	3666	Medium	TWS 3
11	0	0	0	29	435	243	1213	1648	3696	High	5174	High	RTO 1
12	0	0	0	42	630	159	795	1425	2075	High	2214	Low	RTO 1
13	2	0	2	8	120	57	286	408	569	Low	171	Low	RTO 1
14	0	0	0	29	435	190	950	1385	2348	High	4399	High	TWS 1
15	1	7	8	6	90	123	615	713	1028	Low	374	Low	RTO 4
16	0	0	0	5	69	55	274	343	677	Low	3086	Low	RTO 1
17	1	0	1	25	375	252	1260	1636	2773	High	866	Low	RTO 1
18	0	0	0	7	105	335	1675	1780	2671	High	863	Low	RTO 1
19	0	0	0	38	570	214	1069	1639	3587	High	3443	High	TWS 1
20	0	0	0	22	330	167	835	1165	2559	High	2828	High	TWS 2

It was subsequently tested whether statistical significant differences could be found between individuals grouped into high and low MWL-indices when the classification was based on the revised MWL-index (Model 2). Comparisons were made for age, mass, height, BMI, SA, blood pressure heart rate variability variables, smoking or not, length of previous shift, length of test shift, years experience at particular station, shift preferences and time line analyses of the shift when the physiological recordings were made. The results (means of the results over the workshifts and statistical analyses) can be seen in Table 2.3, and the blood pressure arrival values for subjects in Table 2.4.

The experimental data of individuals from the six highest and from the six lowest workload stations, according to the MWL-index (Model 2), were subsequently compared and statistical tests performed to see if they differed significantly. The results can be seen in Table 2.5.

Hereafter the high and low MWL-index groups, according to Model 2 were compared in terms of the way they react to the workload over the total workshifts. The results can be seen in Table 2.6.

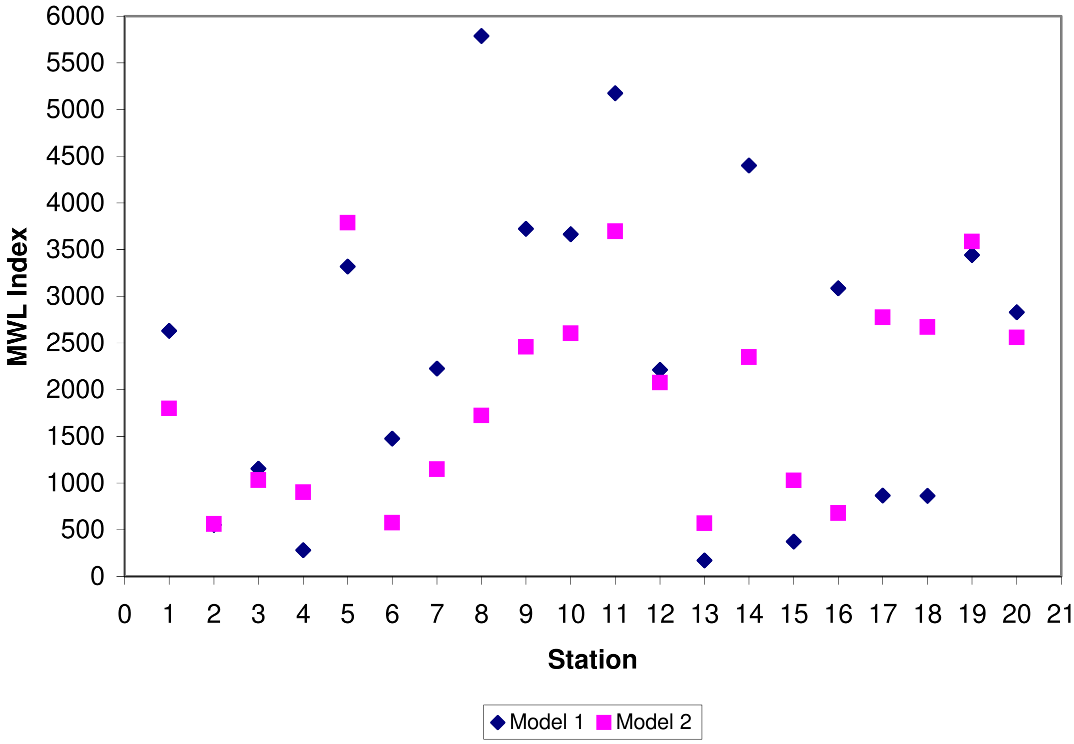


Figure 2.1: Comparison between Model 1 (historical data) and Model 2 (real time data) MWL-index values.

Table 2.3: Mean values over a six hour period (08:00 – 14:00) for subjects divided into high and low stress groups according to Model 2 MWL-index.

Variable	HS	LS	p value
Age	36.7 ± 6.3	43.7 ± 11.5	0.1125
Mass	90.7 ± 29.7	89.7 ± 19.1	0.9325
Height	1.718 ± 0.112	1.738 ± 0.080	0.6441
Body mass index	30.1± 7.4	29.5 ± 5.2	0.8263
Surface area	2.08 ± 0.40	2.09 ± 0.26	0.9533
Cohens	18.7 ± 5.5	20.9 ± 6.4	0.4227
Diastolic blood pressure	83.9 ± 11.7	85.4 ± 9.0	0.7598
Systolic blood pressure	122.1 ± 14.3	122.1 ± 14.1	0.9928
Mean arterial pressure	96.6 ± 12.5	97.6 ± 9.9	0.8472
Pulse pressure	38.2 ± 4.2	36.8 ± 10.0	0.6888
Heart rate (blood pressure monitor)	72.0 ± 8.2	67.8 ± 7.6	0.2665
Cortisol	3.8 ± 1.5	4.2 ± 1.2	0.5643
Heart rate (MiniMitter)	82.6 ± 7.7	77.7 ± 8.6	0.1983
SD of heart rate	5.3 ± 0.9	4.9 ± 1.6	0.5252
RR	741.0 ± 76.0	790.8 ± 93.6	0.2076
SD of RR	41.6 ± 10.4	41.7 ± 17.9	0.9899
RMSSD	29.7 ± 14.0	33.4 ± 15.3	0.5821
SD1	21.3 ± 9.9	23.9 ± 10.9	0.5788
SD2	89.7 ± 19.2	93.4 ± 31.0	0.7511
LF	638.6 ± 269.0	688.4 ± 656.6	0.8280
HF	211.9 ± 207.6	206.8 ± 206.9	0.9567
Ratio (LF/HF)	5.6 ± 4.8	3.9 ± 2.7	0.3244
Total power	850.5 ± 460.4	895.2 ± 843.4	0.8851
Smoking	0.4 ± 0.5	0.5 ± 0.5	0.6733
Length of previous shift	11.6 ± 1.3	10.1 ± 1.9	0.0487
Length of test shift	12.0 ± 0.0	10.5 ± 1.8	No value
Years experience at particular station	4.6 ± 3.3	7.5 ± 7.4	0.2806
Shift preference	1.1 ± 1.0	0.8 ± 0.8	0.4645
Radio communications	166.9 ± 61.9	61.4 ± 30.4	0.0003
Telephone communications	47.4 ± 20.7	28.5 ± 13.9	0.0276
Scheduling	73.2 ± 23.7	26.3 ± 16.3	0.0001
Number of Trains	32.1 ± 22.6	9.0 ± 5.1	0.0105
Number of Authorisations	26.1 ± 10.2	8.8 ± 6.0	0.0002

BMI = Body Mass Index (kg / m^2); SA = Surface Area ($Mass^{0.425}$ x $Height^{0.725}$ x 71.84)

Table 2.4: Mean values at 06:00 for subjects divided into high and low stress group according to Model 2 MWL-index.

Variable	HS	LS	p value
Diastolic	85.9 ± 10.4	89.6 ± 11.2	0.4645
Systolic	125.8 ± 13.5	127.6 ± 16.8	0.2872
MAP	99.2 ± 11.2	102.2 ± 11.8	0.5832
Pulse pressure	39.9 ± 5.6	38.0 ± 13.1	0.6881
Heart rate (blood pressure monitor)	77.1 ± 8.1	76.7 ± 7.5	0.9947
Cortisol	7.7 ± 2.6	7.9 ± 4.2	0.8830

Table 2.5: Selection of the six highest and six lowest MWL-index stations according to the Model 2 MWL-index. Means represent values of subjects grouped into high and low stress groups according to the Model 2 MWL-index. Values represent means over a six hour period (08:00 – 14:00).

Variable	HS	LS	p value
Age	37.3 ± 6.5	44.3 ± 11.1	0.2135
Mass	93.8 ± 34.5	86.2 ± 22.7	0.6612
Height	1.7 ± 0.1	1.7 ± 0.1	0.9521
Body mass index	31.3 ± 8.9	29.3 ± 6.3	0.6599
Surface area	2.11 ± 0.46	2.03 ± 0.31	0.7244
Cohens	18.8 ± 7.0	19.7 ± 8.2	0.8542
Diastolic blood pressure	85.5 ± 9.8	83.9 ± 11.5	0.8188
Systolic blood pressure	122.9 ± 13.8	117.9 ± 14.5	0.5768
Mean arterial pressure	97.9 ± 11.1	95.2 ± 12.4	0.7166
Pulse pressure	37.5 ± 4.2	34.0 ± 4.7	0.2328
Heart rate (blood pressure monitor)	70.9 ± 7.2	71.9 ± 5.9	0.8176
Cortisol	4.1 ± 1.7	4.0 ± 1.3	0.9061
Heart rate (MiniMitter)	82.5 ± 6.5	82.7 ± 6.0	0.9485
SD of heart rate	5.5 ± 1.0	4.9 ± 1.4	0.3605
RR	739.7 ± 54.9	736.3 ± 58.8	0.9214
SD of RR	44.3 ± 12.5	36.1 ± 9.9	0.2376
RMSSD	32.5 ± 15.9	29.8 ± 10.1	0.7307
SD1	23.3 ± 11.2	21.3 ± 7.2	0.7280
SD2	95.8 ± 18.4	82.6 ± 22.1	0.2897
LF	728.4 ± 321.7	430.7 ± 205.7	0.0852
HF	254.7 ± 253.8	159.1 ± 84.1	0.4145
Ratio (LF/HF)	4.9 ± 3.2	3.3 ± 1.5	0.2858
Total power	983.0 ± 560.9	589.8 ± 277.6	0.1547
Smoking	0.5 ± 0.5	0.5 ± 0.5	1.000
Length of previous shift	12.0 ± 0.0	9.0 ± 1.5	
Length of test shift	12.0 ± 0.0	9.7 ± 1.9	
Years experience at particular station	3.4 ± 3.2	9.3 ± 9.1	0.1876
Shift preference	1.2 ± 1.0	1.0 ± 0.9	0.7650
Radio communications	197.8 ± 56.4	46.3 ± 21.5	0.0007
Telephone communications	57.7 ± 19.9	21.8 ± 11.9	0.0036
Scheduling	76.0 ± 23.8	18.5 ± 7.3	0.0013
Number of Trains	34.5 ± 24.7	6.3 ± 2.3	0.0380
Number of Authorisations	25.3 ± 10.5	6.3 ± 2.1	0.0063

Table 2.6: Repeated measures evaluation between the two stress groups (Model 2 MWL-index), change over time for the whole group under investigation and group time interaction for physiological and time line analysis variables.

Variable	Group	Time	Group*Time
Cortisol	0.5643	0.0642	0.4521
Systolic blood pressure	0.0742	0.4159	0.5047
Diastolic blood pressure	0.5784	0.2770	0.7522
Mean arterial pressure	0.2684	0.4305	0.5530
Pulse pressure	0.0122	0.1549	0.9011
Heart rate*	0.2138	0.0001	0.3612
RR*	0.2213	0.0001	0.5683
sdHR*	0.5196	0.0504	0.6269
sdRR*	0.9934	0.0006	0.4212
Body mass index			
Total*	0.8781	0.0060	0.5319
LF power*	0.8260	0.0226	0.5525
LFn	0.4210	0.0703	0.8785
LF peak*	0.2435	0.8383	0.0707
HF power*	0.9762	0.0056	0.4778
HFn	0.4210	0.0703	0.8785
HF peak*	0.1601	0.4830	0.0536
Ratio (LF/HF)*	0.2863	0.0866	0.8189
RMSSD*	0.5784	0.0017	0.5659
SD1*	0.5753	0.0017	0.5619
SD1n*	0.7424	0.0078	0.5472
SD2*	0.8046	0.0017	0.4962
SD2n*	0.7253	0.1413	0.7272
Number of authorisations*	0.0002	0.2093	0.2001
Radio communications*	0.0001	0.4620	0.5608
Scheduling*	0.0001	0.4975	0.4418
Telephone communications*	0.0350	0.1424	0.1442
Number of trains*	0.0064	0.5394	0.0453

* = Hour intervals

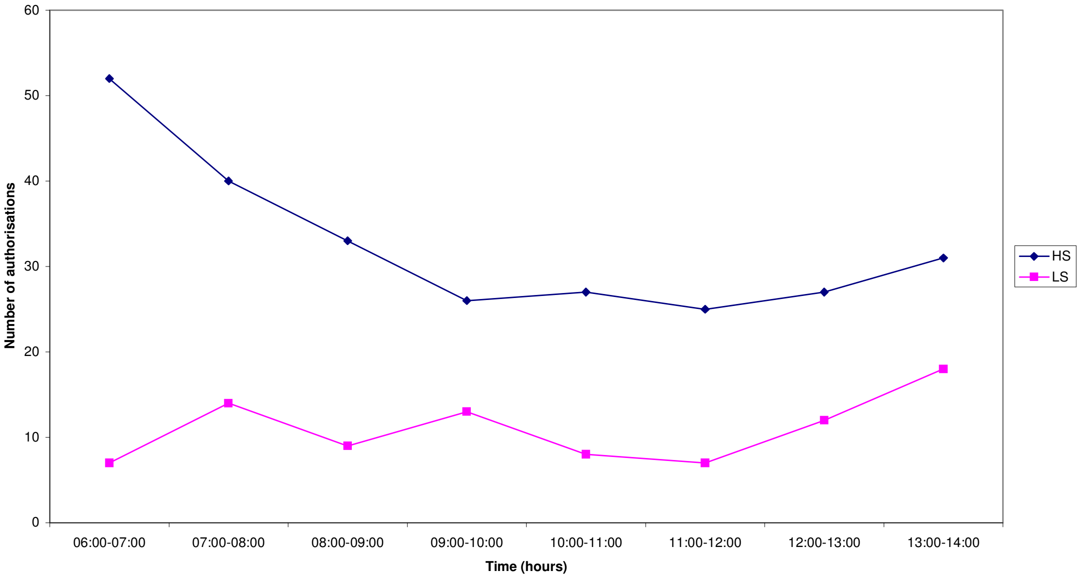


Figure 2.2: Mean total number of authorizations for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

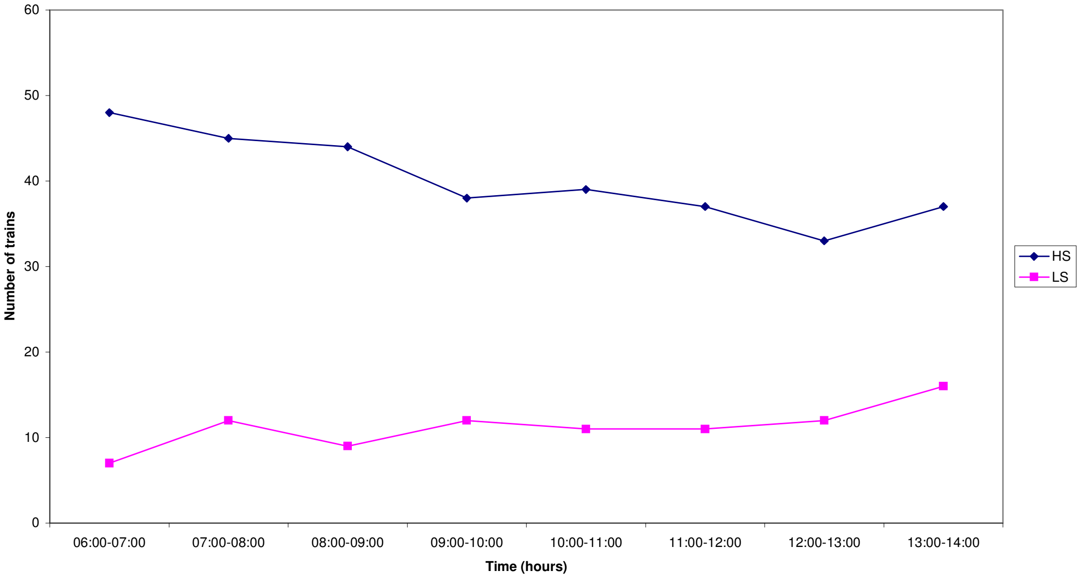


Figure 2.3: Mean total number of trains for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

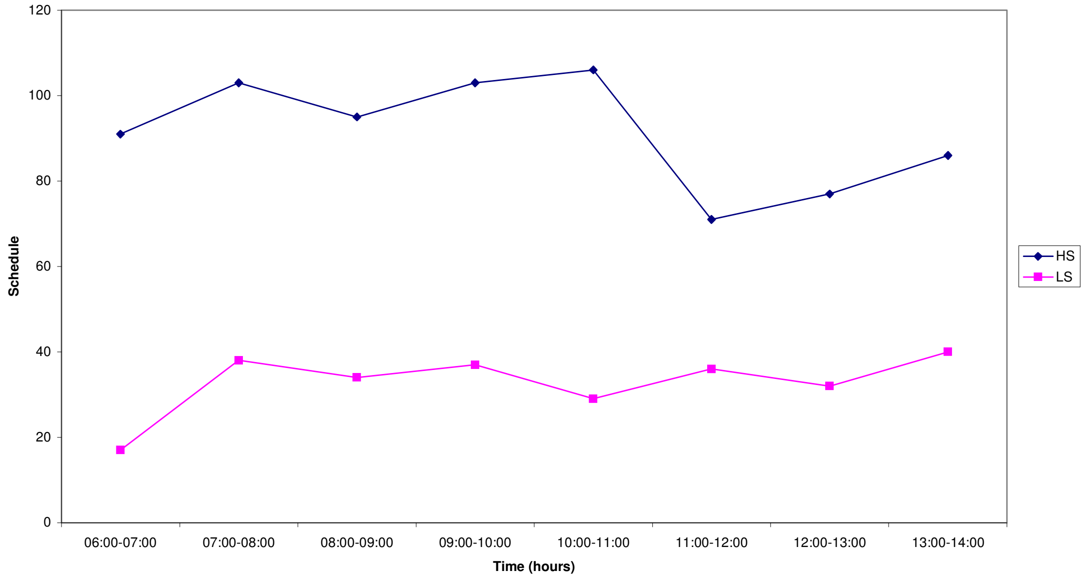


Figure 2.4: Mean total number of schedule transactions for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

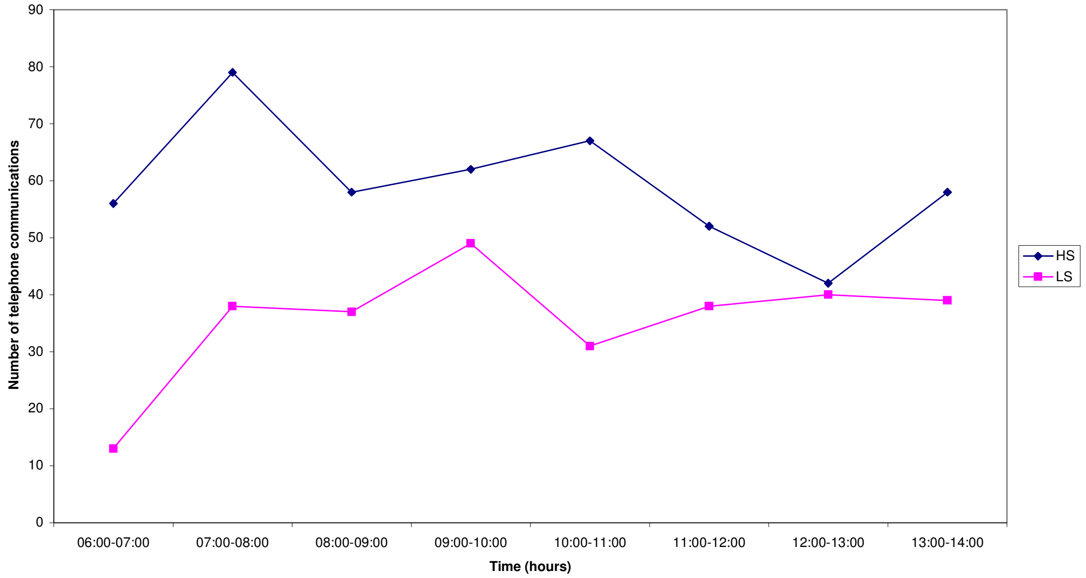


Figure 2.5: Mean total number of telephone communications for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

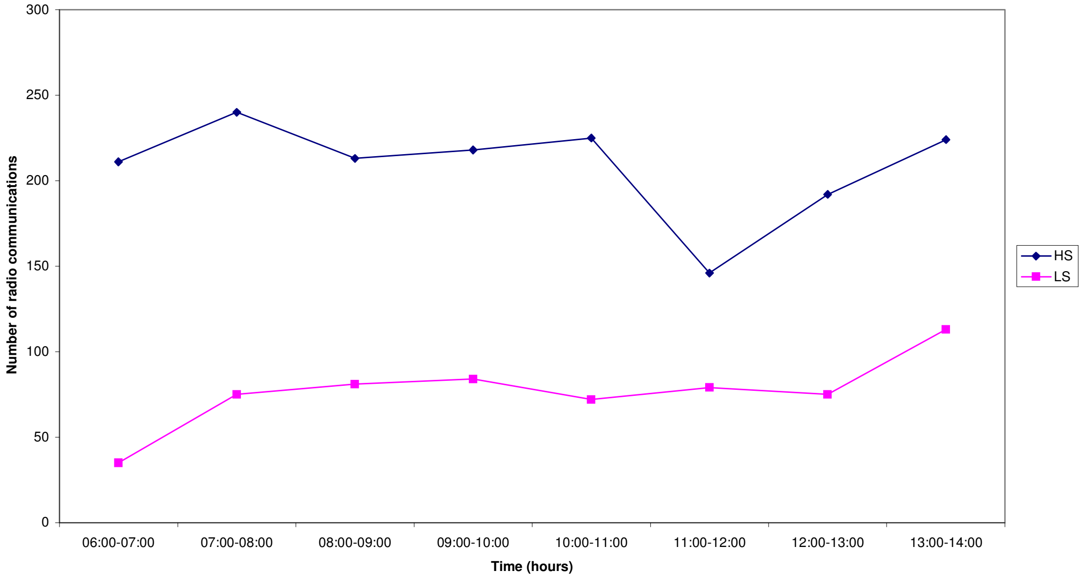


Figure 2.6: Mean total number of radio communications for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

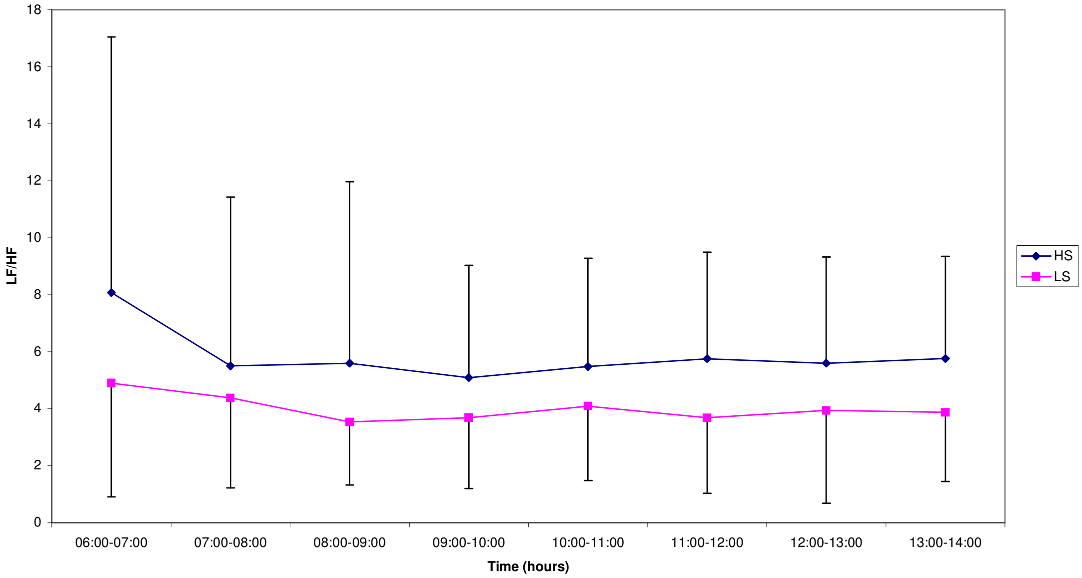


Figure 2.7: Mean ratio (LF/HF) for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

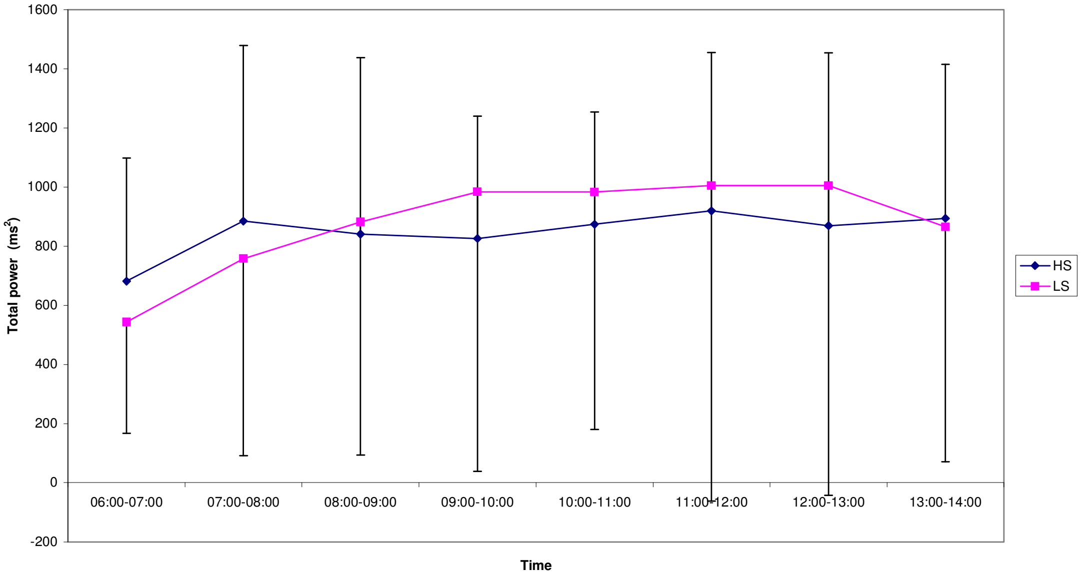


Figure 2.8: Mean total power for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

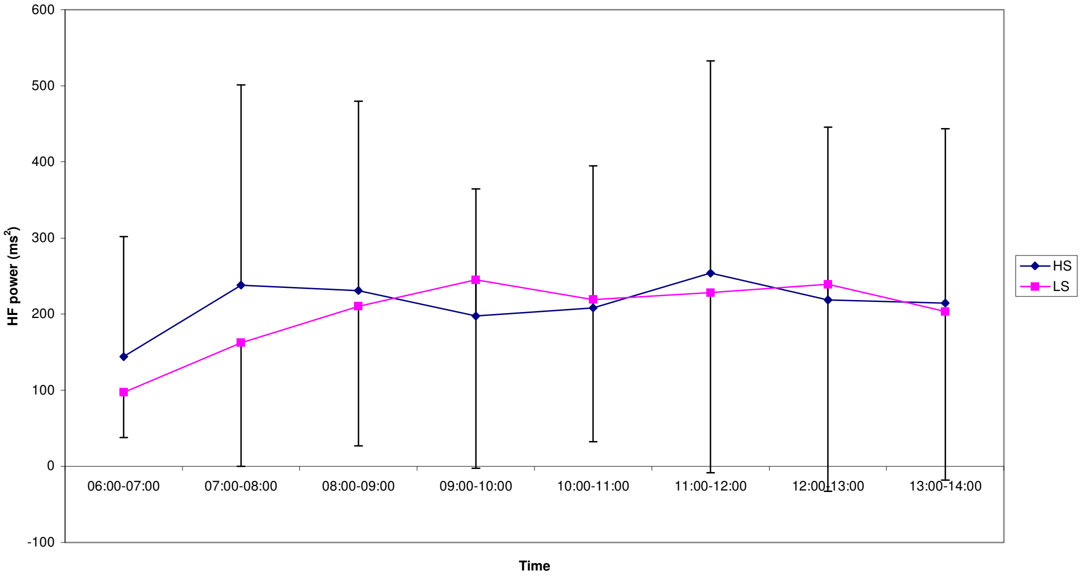


Figure 2.9: Mean high frequency (HF) power for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

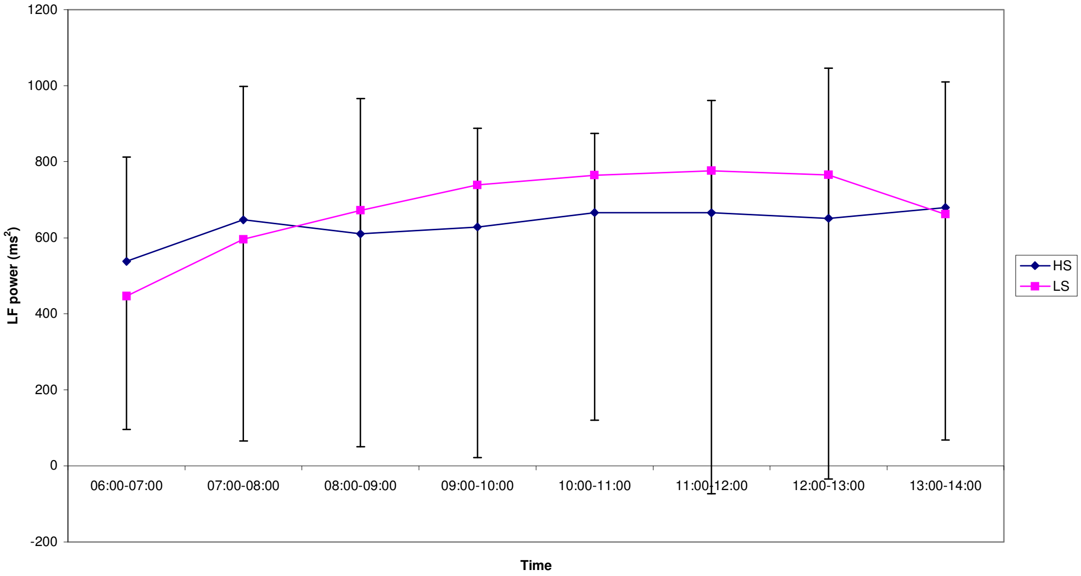


Figure 2.10: Mean low frequency power (LF) for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

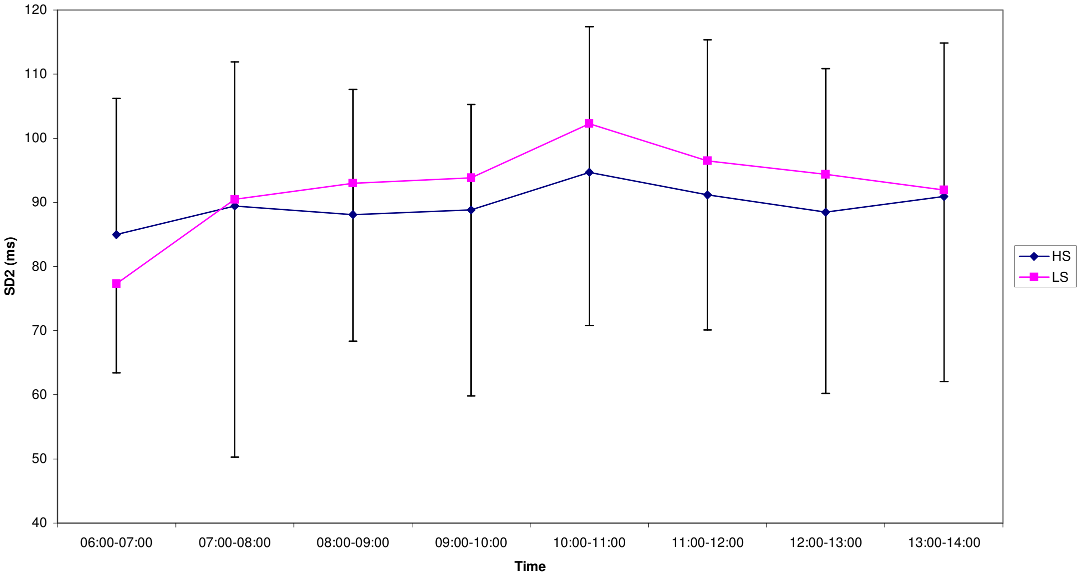


Figure 2.11: Mean SD2 for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

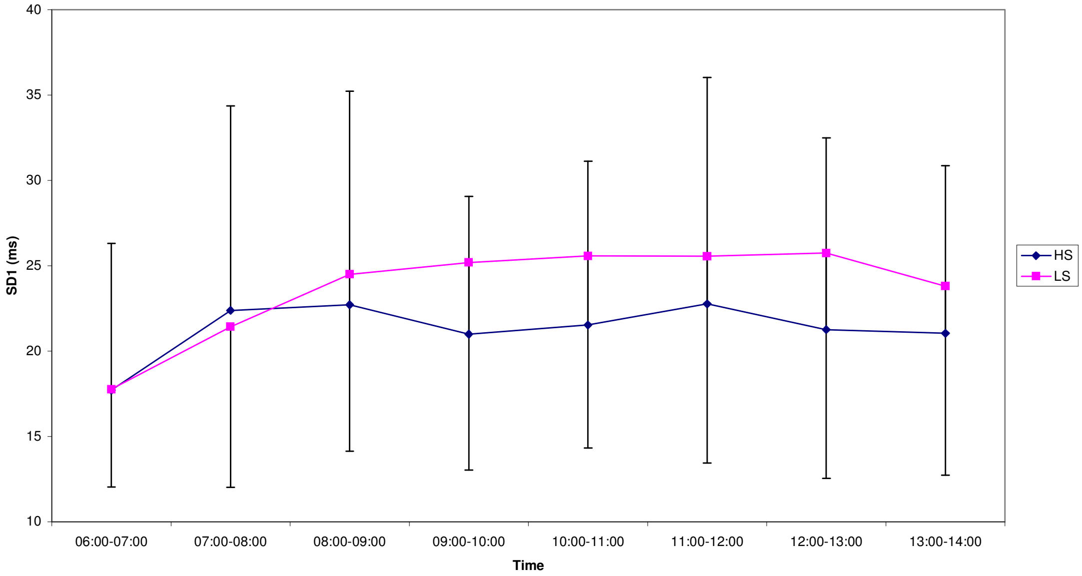


Figure 2.12: Mean SD1 for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

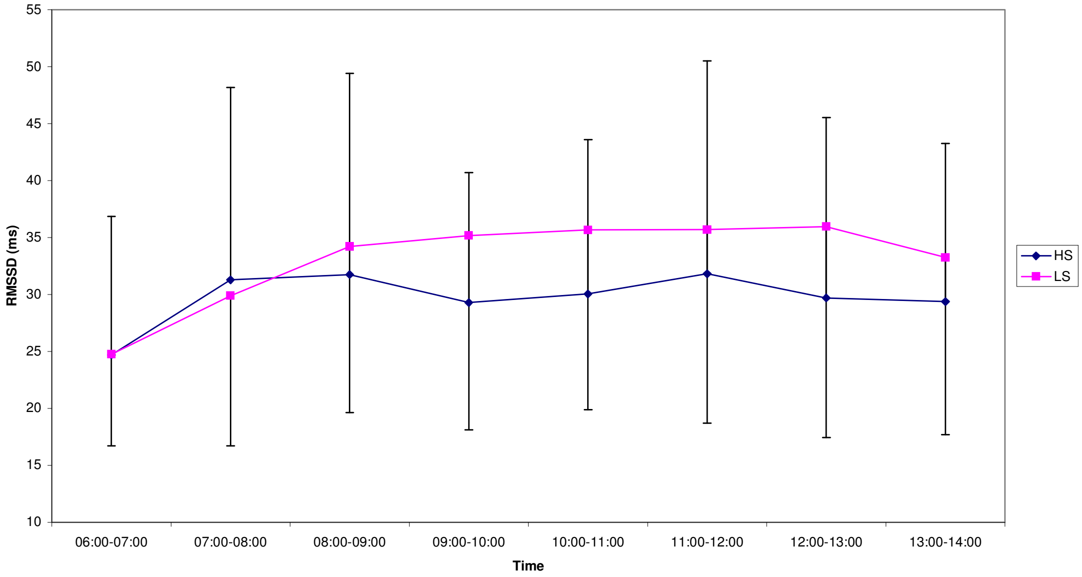


Figure 2.13: Mean RMSSD for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

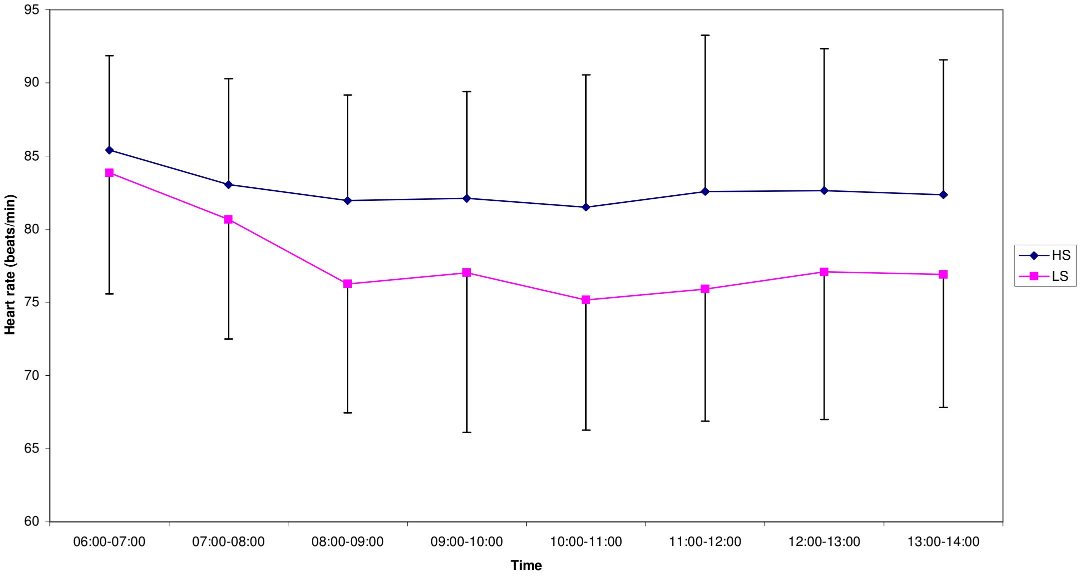


Figure 2.14: Mean heart rate (Minimitter) for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

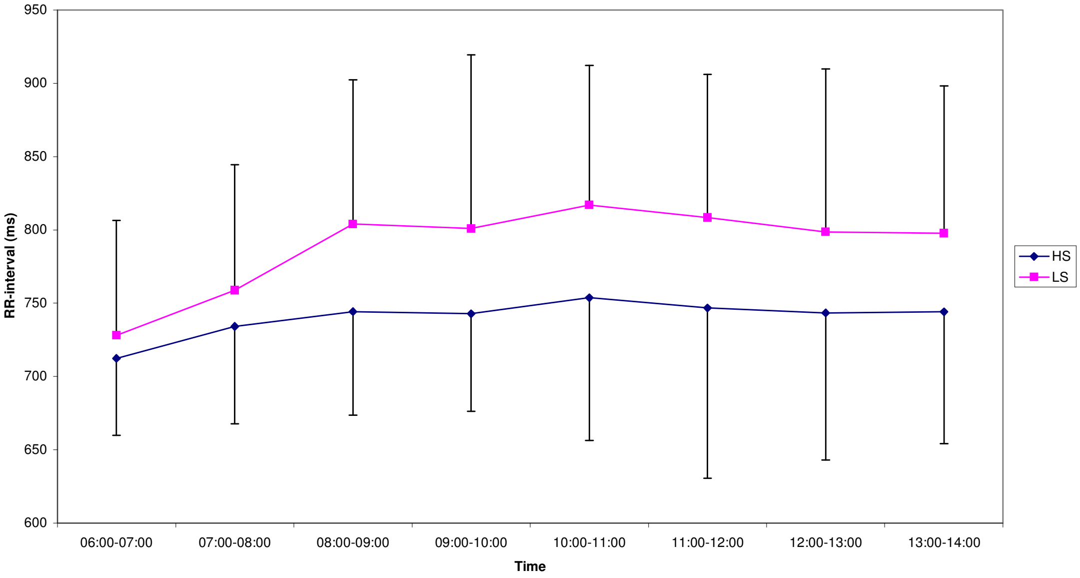


Figure 2.15: Mean RR-interval for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

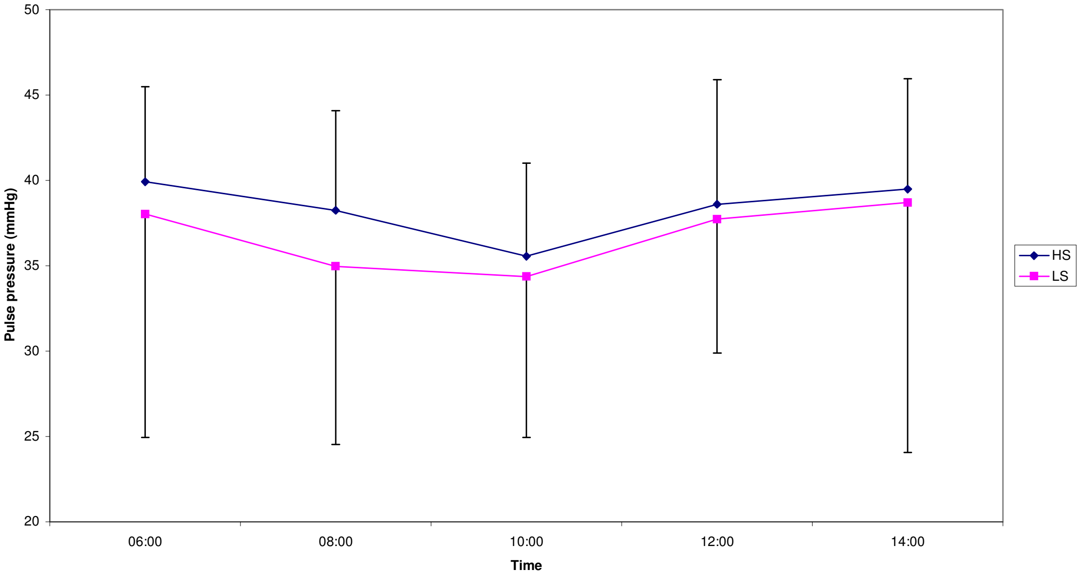


Figure 2.16: Mean pulse pressure for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

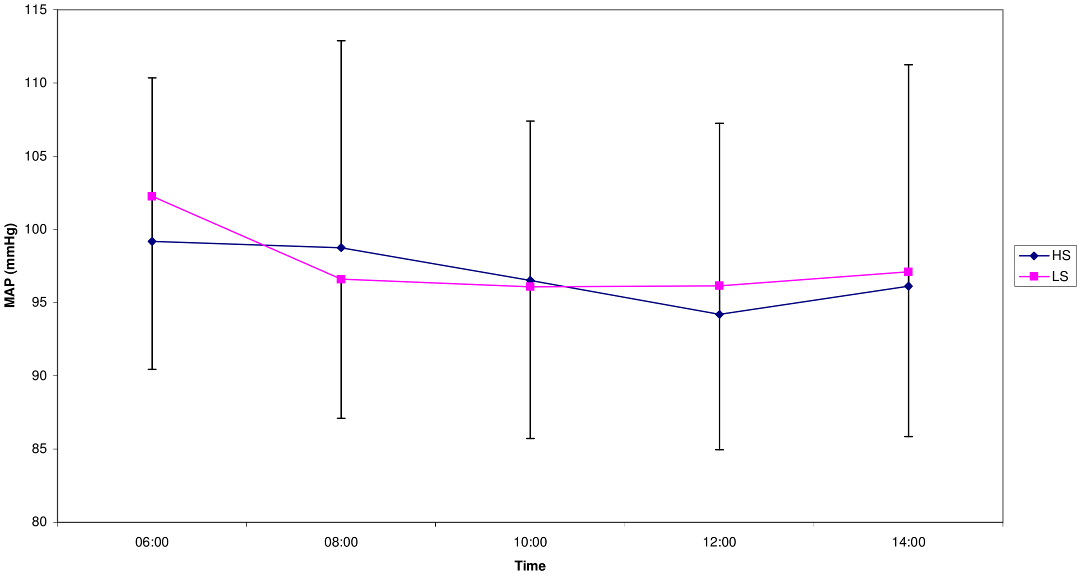


Figure 2.17: Mean arterial pulse pressure (MAP) for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

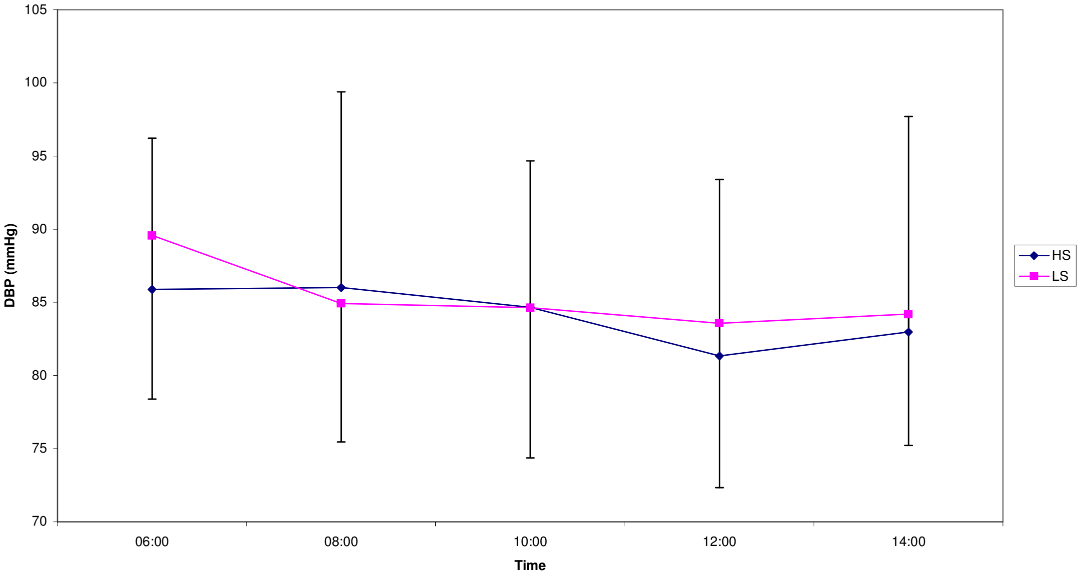


Figure 2.18: Mean diastolic blood pressure (DBP) for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

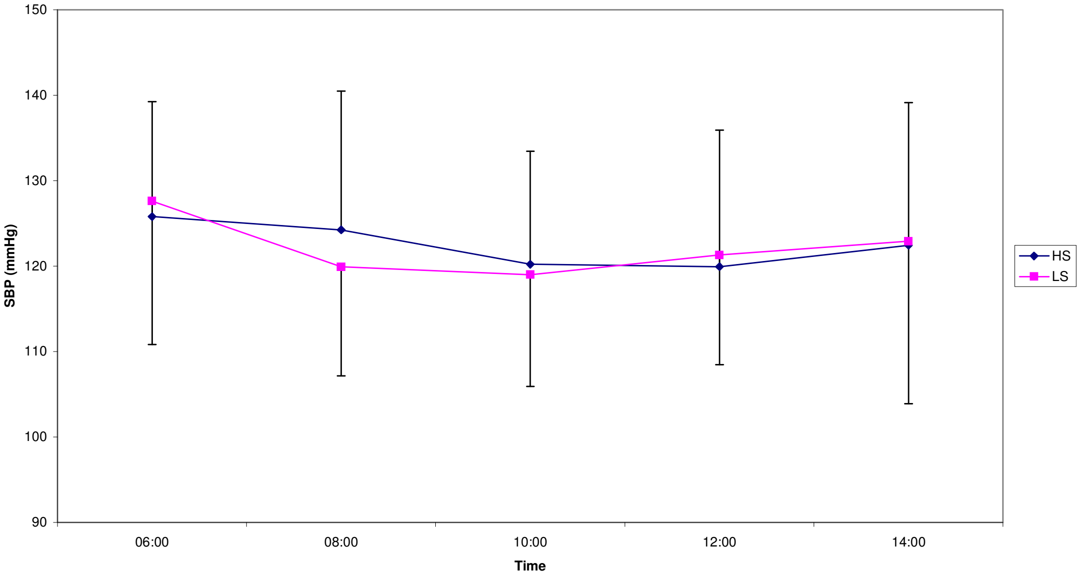


Figure 2.19: Mean systolic blood pressure (SBP) for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

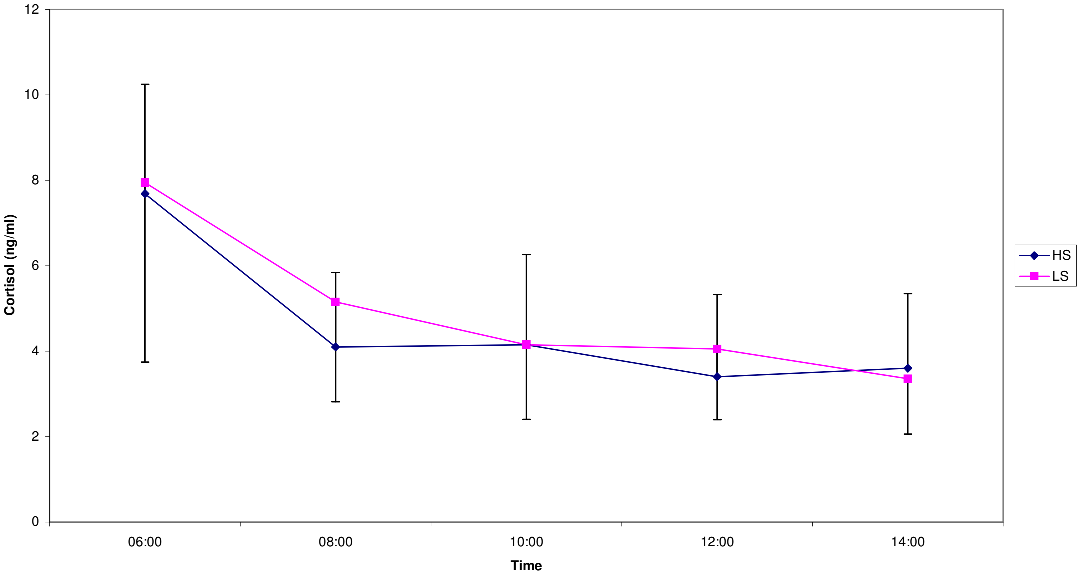


Figure 2.20: Mean cortisol for TCOs designated to fall into the high (HS) or low (LS) stress group according to the MWL-index (Model 2).

3.3 Physiological Parameters: Separation of Work Stations Into High and Low Stress as Extrapolated from the Stress Levels of the Individuals at the Work Stations

The analyses here was performed in terms of an adapted allostatic load measurement, only those parameters that could be measured within the possibilities of the study could be included. Those measurements that required venipuncture could not be performed. The data used in the calculation of the numerical value for each individual factor or parameter of the allostatic load were derived from the mean of the values obtained over the work shift. The reason for this was that baseline values in this group could not really be seen as base line, as all of them appeared to arrive with values higher than any reached during the experimental procedure. The fact that the stressor, in this case is not a once off stressor, but that the TCOs live with this level of stress axes activity for the major part of everyday, supports the decision to use the mean values over the work shift. The values of the following parameters were included: salivary cortisol and BMI as an indices of hypothalamo-pituitary-adrenocortical axis activity, systolic and diastolic pressure as measure, of cardiovascular activity - largely reflecting sympatho-adrenomedullary axis activation, heart rate as an indicator of sympatho-adrenomedullary activation and heart rate variability variables as indicators of autonomic activity. The values of each individual for each of the 5 indicators were classified according to the 50th percentile. Allostatic load was then calculated by summing up the number of parameters for which the subject fell into the highest risk, i.e., above the 50 percentile. Subjects were subsequently ranked according to their total hits.

Table 3.1 shows combinations of factors used to calculate the adapted allostatic load and classifications into high and low groups according to the values of different combinations of physiological stress indicators. The decision whether an individual was classified with a high or low allostatic load depended on whether the sum was greater or equal to 3 (high) or not (low). Three models, i.e., A, B and C were developed – depending on the combination of factors included. The significance of the statistical differences for the three models, i.e., the three combinations of factors can be seen in Tables 3.2, 3.3 and 3.4.

The physiological rating of the three models can be seen in Tables 3.5, 3.6 and 3.7. Graphical presentations of the time line variables and the physiological measures can be seen in Figures 3.1 to 3.19.

Table 3.1 Calculation of allostatic load index

	Physiological variables measure that are associated with allostatic load						50th Percentile						Allostatic load index			Allostatic group		
Id	BMI	Cortisol	Systolic	Diastolic	HR	LF	BMI	Cortisol	Systolic	Diastolic	HR	LF	A,B,C,D,E,F (A)	A,B,C,D,E (B)	B,C,D,E (C)	A,B,C,D,E,F (A)	A,B,C,D,E (B)	B,C,D,E (C)
	A	B	C	D	E	F	28.4	3.4	119.3	84.1	81.8	553.1						
1	27.7	3.6	118.9	85.9	76.6	1262.4		1		1			2	2	2	Low	Low	Low
2	20.7	2.5	93.1	66.4	87.5	430.0					1	1	2	1	1	Low	Low	Low
3	27.8	3.3	119.3	86.7	66.1	2296.0				1			1	1	1	Low	Low	Low
4	40.2	3.3	130.6	97.3	88.1	140.9	1		1	1	1	1	5	4	3	High	High	High
5	36.1	3.0	137.8	96.2	81.5	597.5	1		1	1			3	3	2	High	High	Low
6	28.1	3.4	116.8	83.5	82.1	727.9					1		1	1	1	Low	Low	Low
7	28.5	4.0	134.2	94.1	84.2	583.9	1	1	1	1	1		5	5	4	High	High	High
8	28.4	5.0	124.5	89.9	74.1	648.3	1	1	1	1			4	4	3	High	High	High
9	30.2	3.3	119.8	81.5	88.5	479.1	1		1		1	1	4	3	2	High	High	Low
10	28.5	5.0	114.3	78.3	77.1	1030.4	1	1					2	2	1	Low	Low	Low
11	46.8	7.1	*	*	77.9	456.6	1	1	1	1		1	5	4	3	High	High	High
12	34.1	1.9	143.8	98.3	89.7	584.9	1		1	1	1		4	4	3	High	High	High
13	27.3	4.8	113.6	75.8	82.8	312.3		1			1	1	3	2	2	High	Low	Low
14	26.8	5.3	117.4	84.2	85.3	429.1		1		1	1	1	4	3	3	High	High	High
15	35.1	5.9	143.8	82.3	63.8	93.2	1	1	1			1	4	3	2	High	High	Low
16	31.0	6.1	113.1	81.5	71.5	388.9	1	1				1	3	2	1	High	Low	Low
17	28.3	2.9	121.4	84.1	80.9	1167.7			1				1	1	1	Low	Low	Low
18	26.6	3.1	134.3	96.2	95.0	353.2			1	1	1	1	4	3	3	High	High	High
19	21.5	3.5	105.8	73.9	82.4	764.8		1			1		2	2	2	Low	Low	Low
20	22.1	3.1	100.8	61.1	67.4	522.3						1	1	0	0	Low	Low	Low
							10	10	10	10	10	10						

* - it was assumed that this subject had high blood pressure

Table 3.2: Mean values over a six hour period (08:00 – 14:00) for subjects divided into high and low stress groups according to Model A of allostatic load.

Variable	HS	LS	p value
Age	45.3 ± 8.0	32.5 ± 6.5	0.0013
Mass	100.0 ± 22.8	75.5 ± 19.5	0.0225
Height	1.7 ± 0.1	1.7 ± 0.1	0.3639
Body mass index	32.6 ± 6.2	25.6 ± 3.5	0.0047
Surface area	2.21 ± 0.28	1.90 ± 0.32	0.0291
Cohens	20.3 ± 5.9	19.1 ± 6.4	0.6899
Diastolic blood pressure	89.5 ± 7.3	78.0 ± 9.8	0.0091
Systolic blood pressure	129.7 ± 10.8	111.6 ± 10.3	0.0019
Mean arterial pressure	102.9 ± 7.6	89.2 ± 9.8	0.0031
Pulse pressure	40.2 ± 8.6	33.6 ± 4.1	0.0431
Heart rate (blood pressure monitor)	72.9 ± 7.5	65.5 ± 6.8	0.0415
Cortisol	4.4 ± 1.6	3.4 ± 0.7	0.0766
Heart rate (MiniMitter)	81.9 ± 8.8	77.5 ± 7.5	0.2655
SD of heart rate	4.5 ± 1.1	6.1 ± 0.9	0.0024
RR	748.5 ± 88.1	792.0 ± 83.6	0.2851
SD of RR	33.6 ± 7.6	53.8 ± 13.4	0.0031
RMSSD	23.8 ± 8.4	43.2 ± 13.9	0.0010
SD1	17.1 ± 5.9	30.9 ± 9.9	0.0010
SD2	79.1 ± 17.5	110.1 ± 24.2	0.0037
LF	422.3 ± 176.9	1025.2 ± 592.2	0.0238
HF	105.4 ± 65.1	365.3 ± 239.9	0.0181
Ratio (LF/HF)	5.5 ± 4.3	3.6 ± 3.0	0.2752
Total power	527.7 ± 219.5	1390.5 ± 776.2	0.0161
Smoking	0.4 ± 0.5	0.5 ± 0.5	0.7309
Length of previous shift	10.8 ± 1.7	10.9 ± 1.8	0.9595
Length of test shift	11.5 ± 1.2	10.9 ± 1.8	0.3577
Years experience at particular station	6.8 ± 6.6	4.8 ± 4.5	0.4689
Shift preference	0.9 ± 0.9	1.0 ± 0.9	0.8433
Radio communications	117.6 ± 81.6	109.0 ± 59.4	0.8018
Telephone communications	37.0 ± 21.8	39.4 ± 17.4	0.7998
Scheduling	51.1 ± 33.1	47.8 ± 30.1	0.8219
Number of Trains	22.8 ± 24.1	17.1 ± 11.8	0.4912
Number of Authorisations	16.3 ± 12.0	19.2 ± 12.6	0.5989

Table 3.3: Mean values over a six hour period (08:00 – 14:00) for subjects divided into high and low stress groups according to Model B of allostatic load.

Variable	HS	LS	p value
Age	43.8 ± 7.8	36.6 ± 10.4	0.0970
Mass	102.0 ± 24.6	78.5 ± 18.3	0.0261
Height	1.7 ± 0.1	1.7 ± 0.1	0.04896
Body mass index	33.3 ± 6.5	26.3 ± 3.5	0.0102
Surface area	2.23 ± 0.31	1.94 ± 0.29	0.0423
Cohens	21.5 ± 5.5	18.1 ± 6.2	0.2092
Diastolic blood pressure	91.5 ± 6.5	78.6 ± 8.9	0.0023
Systolic blood pressure	132.6 ± 9.8	112.7 ± 9.3	0.0003
Mean arterial pressure	105.2 ± 6.4	89.9 ± 8.8	0.0005
Pulse pressure	41.1 ± 9.3	34.1 ± 3.9	0.0608
Heart rate (blood pressure monitor)	74.0 ± 7.3	66.0 ± 6.7	0.0230
Cortisol	4.2 ± 1.6	3.8 ± 1.1	0.05643
Heart rate (MiniMitter)	82.8 ± 9.0	77.5 ± 7.1	0.1566
SD of heart rate	4.5 ± 1.2	5.7 ± 1.2	0.0365
RR	740.5 ± 91.0	791.3 ± 78.7	0.1987
SD of RR	33.4 ± 8.4	49.8 ± 14.4	0.0061
RMSSD	22.6 ± 8.3	40.6 ± 13.7	0.0023
SD1	16.2 ± 5.9	29.0 ± 9.7	0.0022
SD2	78.9 ± 18.9	104.1 ± 25.1	0.0208
LF	436.7 ± 191.2	890.3 ± 595.0	0.0427
HF	100.3 ± 68.1	318.4 ± 234.3	0.0172
Ratio (LF/HF)	6.0 ± 4.6	3.5 ± 2.7	0.1419
Total power	537.0 ± 241.5	1208.7 ± 784.6	0.0258
Smoking	0.4 ± 0.5	0.5 ± 0.5	0.6733
Length of previous shift	10.8 ± 1.9	10.9 ± 1.6	0.9010
Length of test shift	11.6 ± 1.3	10.9 ± 1.6	0.2912
Years experience at particular station	5.7 ± 3.8	6.4 ± 7.5	0.7961
Shift preference	0.9 ± 0.9	1.0 ± 0.9	0.8086
Radio communications	135.5 ± 77.3	92.8 ± 62.9	0.1919
Telephone communications	40.2 ± 21.7	35.7 ± 18.4	0.6230
Scheduling	58.2 ± 31.6	41.3 ± 29.9	0.2349
Number of Trains	26.1 ± 25.3	15.0 ± 11.3	0.2280
Number of Authorisations	18.4 ± 12.1	16.5 ± 12.6	0.7344

Table 3.4: Mean values over a six hour period (08:00 – 14:00) for subjects divided into high and low stress groups according to Model C of allostatic load.

Variable	HS	LS	p value
Age	43.4 ± 6.0	38.5 ± 11.0	0.2857
Mass	100.3 ± 27.9	84.8 ± 21.3	0.1800
Height	1.7 ± 0.1	1.7 ± 0.1	0.8335
Body mass index	33.1 ± 7.8	28.0 ± 4.6	0.0849
Surface area	2.21 ± 0.34	2.02 ± 0.32	0.2417
Cohens	22.1 ± 5.8	18.5 ± 5.8	0.2039
Diastolic blood pressure	93.7 ± 5.4	80.5 ± 9.0	0.0044
Systolic blood pressure	131.4 ± 8.7	117.8 ± 13.8	0.0429
Mean arterial pressure	106.3 ± 6.3	93.0 ± 10.0	0.0087
Pulse pressure	37.7 ± 4.7	37.3 ± 8.8	0.09240
Heart rate (blood pressure monitor)	76.3 ± 5.4	66.8 ± 7.2	0.0108
Cortisol	4.2 ± 1.7	3.9 ± 1.2	0.5781
Heart rate (MiniMitter)	84.9 ± 7.1	77.6 ± 8.1	0.0590
SD of heart rate	4.7 ± 1.0	5.3 ± 1.4	0.3086
RR	719.2 ± 65.7	791.1 ± 88.5	0.0767
SD of RR	34.1 ± 8.6	45.7 ± 15.3	0.0834
RMSSD	22.4 ± 10.2	36.5 ± 14.2	0.0320
SD1	16.1 ± 7.2	26.1 ± 10.1	0.0320
SD2	79.5 ± 22.6	98.0 ± 24.9	0.1183
LF	456.7 ± 173.5	774.8 ± 570.3	0.0822
HF	101.7 ± 79.8	267.3 ± 225.6	0.0295
Ratio (LF/HF)	6.9 ± 5.1	3.6 ± 2.6	0.1442
Total power	558.4 ± 241.6	1042.1 ± 758.6	0.0512
Smoking	0.6 ± 0.5	0.4 ± 0.5	0.4499
Length of previous shift	10.3 ± 2.1	11.2 ± 1.5	0.2953
Length of test shift	11.4 ± 1.5	11.2 ± 1.5	0.6967
Years experience at particular station	5.6 ± 4.2	6.2 ± 6.7	0.8155
Shift preference	0.6 ± 0.8	1.2 ± 0.9	0.1672
Radio communications	138.1 ± 79.9	101.2 ± 67.0	0.2858
Telephone communications	38.0 ± 20.3	37.9 ± 20.2	0.9936
Scheduling	64.7 ± 35.5	41.7 ± 26.6	0.1173
Number of Trains	30.7 ± 29.5	15.1 ± 9.9	0.2169
Number of Authorisations	20.0 ± 13.7	16.1 ± 11.4	0.5020

Table 3.5 Physiological rating, Model A: Cortisol, systolic blood pressure, diastolic blood pressure, heart rate, LF power and BMI

Variable	Group	Time	Group*Time
Cortisol	0.3879	0.1357	0.4345
Systolic blood pressure	0.0320	0.6155	0.9046
Diastolic blood pressure	0.4240	0.2186	0.4327
Mean arterial pressure	0.0296	0.4726	0.5975
Pulse pressure	0.4000	0.1775	0.8002
Heart rate*	0.2193	0.0002	0.2067
RR*	0.1743	0.0002	0.1564
sdHR*	0.9401	0.0592	0.4809
sdRR*	0.4943	0.0003	0.0521
Body mass index			
Total*	0.6556	0.0010	0.0278
LF power*	0.7133	0.0110	0.2162
LFn	0.9104	0.0299	0.0008
LF peak*	0.5627	0.6572	0.1180
HF power*	0.5980	0.0001	0.0003
HFn	0.9104	0.0299	0.0008
HF peak*	0.9682	0.5256	0.3921
Ratio (LF/HF)*	0.9907	0.0910	0.8633
RMSSD*	0.7347	0.0002	0.0046
SD1*	0.7310	0.0002	0.0047
SD1n*	0.8134	0.0015	0.0058
SD2*	0.2146	0.0041	0.0987
SD2n*	0.4607	0.2223	0.3031
Number of authorisations*	0.4821	0.2192	0.8366
Radio communications*	0.3090	0.4045	0.2866
Scheduling*	0.3314	0.4724	0.3623
Telephone communications*	0.7846	0.0145	0.1530
Number of trains*	0.6477	0.4917	0.9929

* = Hour intervals

Table 3.6 Physiological rating, Model B: Cortisol, systolic blood pressure, diastolic blood pressure, heart rate and BMI

Variable	Group	Time	Group*Time
Cortisol	0.5643	0.0691	0.7287
Systolic blood pressure	0.0104	0.3500	0.1504
Diastolic blood pressure	0.4544	0.3248	0.2275
Mean arterial pressure	0.1123	0.4758	0.1613
Pulse pressure	0.0057	0.1210	0.3002
Heart rate*	0.1711	0.0001	0.2525
RR*	0.2131	0.0002	0.1923
sdHR*	0.0422	0.0410	0.4505
sdRR*	0.0067	0.0004	0.1138
Body mass index			
Total*	0.0209	0.0026	0.0852
LF power*	0.0368	0.0183	0.3780
LFn	0.1018	0.0153	0.0001
LF peak*	0.4949	0.8378	0.1677
HF power*	0.0144	0.0006	0.0018
HFn	0.1018	0.0153	0.0001
HF peak*	0.9073	0.4853	0.4328
Ratio (LF/HF)*	0.1251	0.0837	0.8213
RMSSD*	0.0027	0.0006	0.0164
SD1*	0.0026	0.0006	0.0170
SD1n*	0.0026	0.0036	0.0245
SD2*	0.0204	0.0030	0.3261
SD2n*	0.0328	0.0911	0.6028
Number of authorisations*	0.7920	0.1578	0.7340
Radio communications*	0.1926	0.4240	0.1506
Scheduling*	0.2347	0.4533	0.1058
Telephone communications*	0.6673	0.1389	0.1672
Number of trains*	0.2401	0.4861	0.9473

* = Hour intervals

Table 3.7: Physiological rating, Model C: Cortisol, systolic blood pressure, diastolic blood pressure and heart rate

Variable	Group	Time	Group*Time
Cortisol	0.3498	0.1352	0.5378
Systolic blood pressure	0.1667	0.5201	0.8207
Diastolic blood pressure	0.1059	0.7820	0.1343
Mean arterial pressure	0.1116	0.8103	0.2841
Pulse pressure	0.6299	0.2645	0.7071
Heart rate*	0.1178	0.0001	0.2759
RR*	0.2059	0.0002	0.5233
sdHR*	0.6746	0.0596	0.8012
sdRR*	0.2260	0.0011	0.4261
Body mass index			
Total*	0.2983	0.0146	0.4374
LF power*	0.3212	0.0489	0.7297
LFn	0.2893	0.0359	0.1760
LF peak*	0.3055	0.9320	0.1667
HF power*	0.3234	0.0100	0.1199
HFn	0.2893	0.0359	0.1760
HF peak*	0.9900	0.5953	0.8676
Ratio (LF/HF)*	0.1695	0.0149	0.2808
RMSSD*	0.1659	0.0019	0.1538
SD1*	0.1663	0.0019	0.1549
SD1n*	0.1848	0.0085	0.1667
SD2*	0.2360	0.0014	0.4467
SD2n*	0.3766	0.1078	0.5326
Number of authorisations*	0.1690	0.1190	0.9517
Radio communications*	0.2241	0.3864	0.2522
Scheduling*	0.3541	0.3858	0.3335
Telephone communications*	0.7592	0.1702	0.1385
Number of trains*	0.0451	0.6096	0.7609

* = Hour intervals

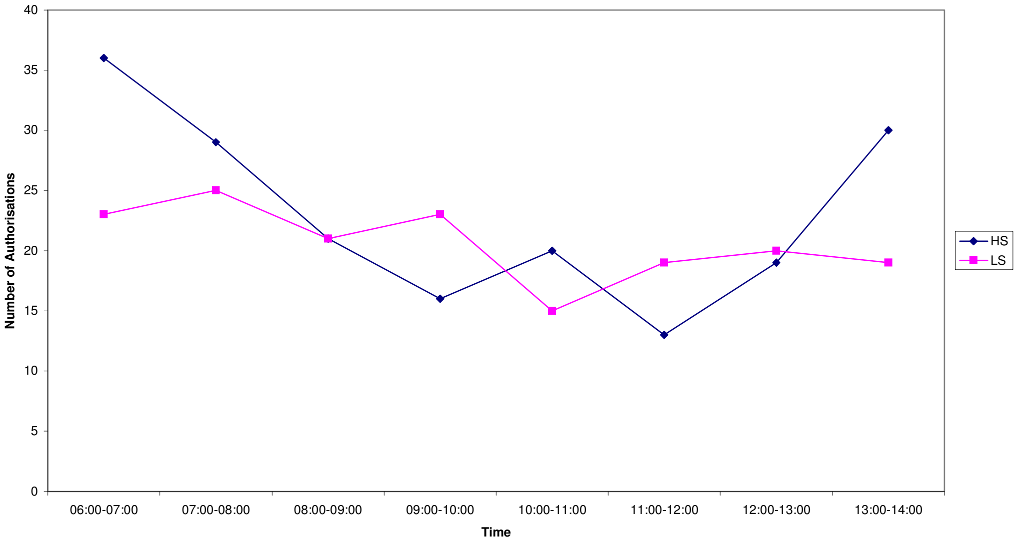


Figure 3.1: Mean total number of authorizations for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

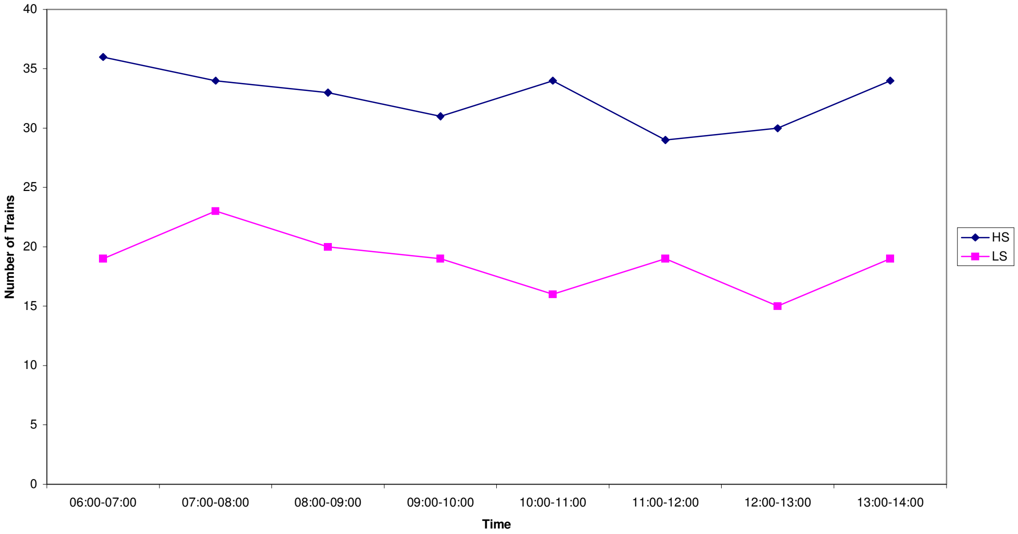


Figure 3.2: Mean total number of trains for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

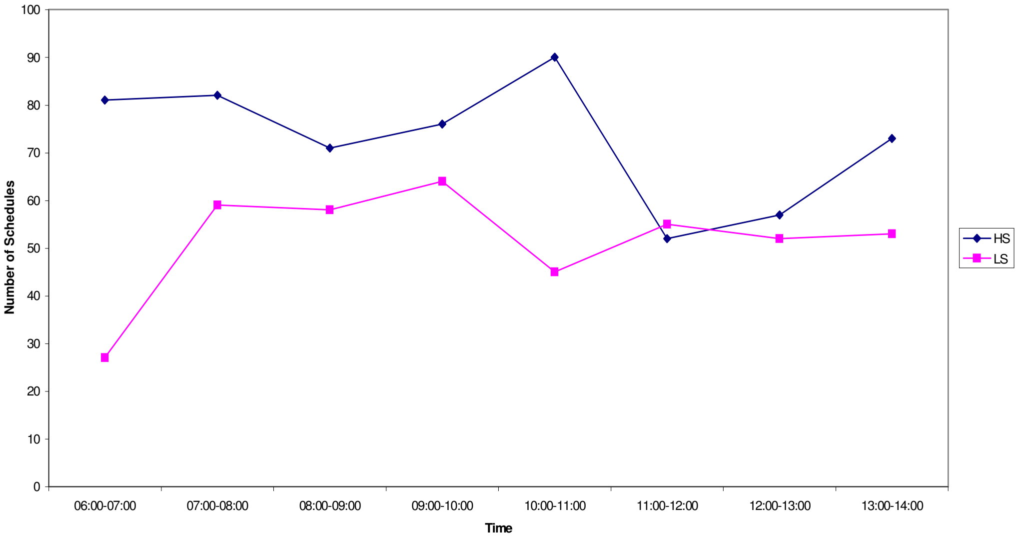


Figure 3.3: Mean total number of schedule transactions for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

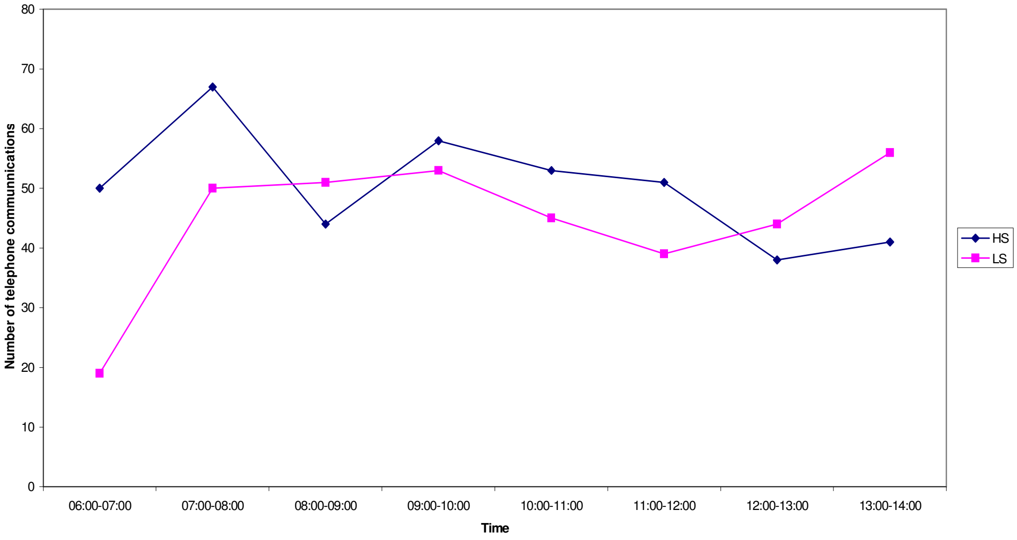


Figure 3.4: Mean total number of telephone communications for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

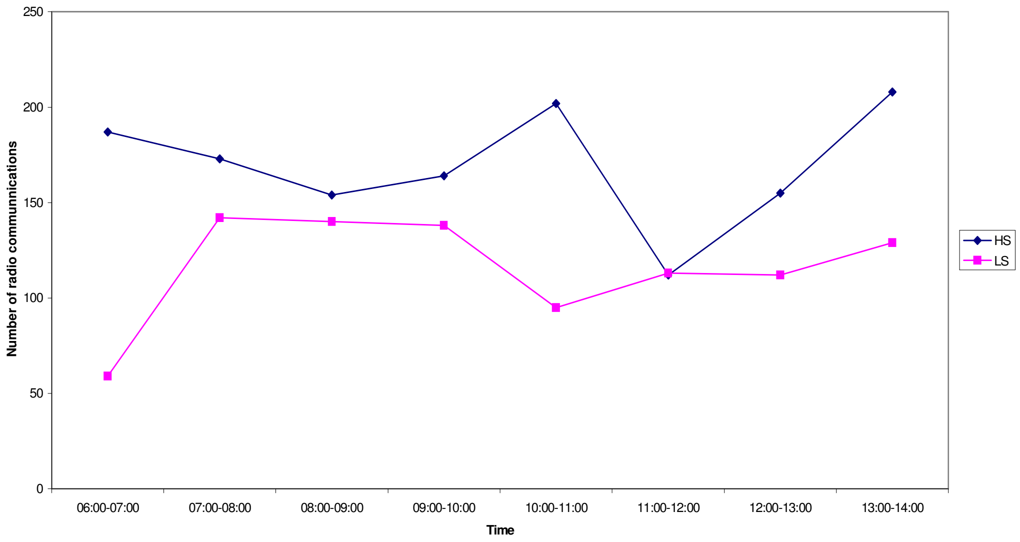


Figure 3.5: Mean total number of radio communications for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

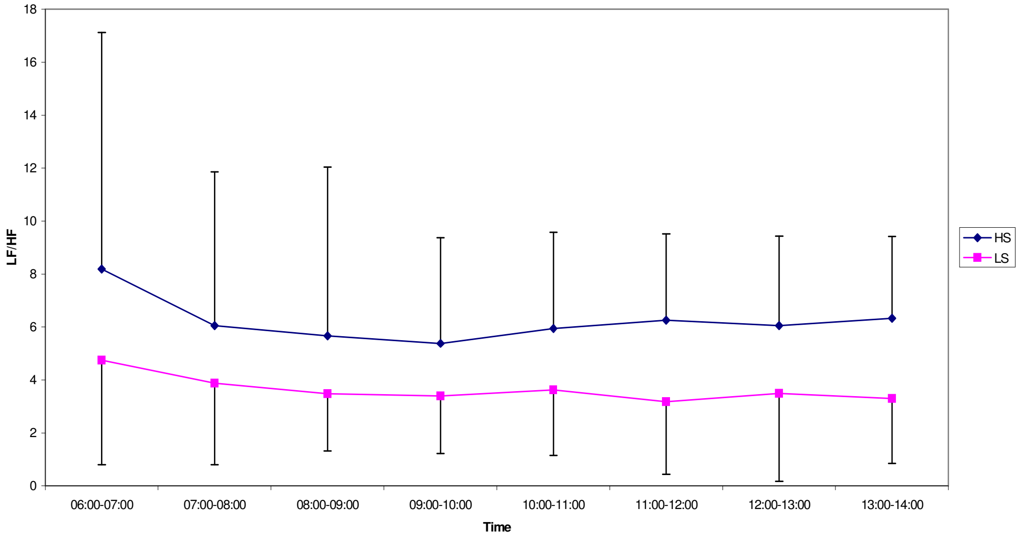


Figure 3.6: Mean ratio (LF/HF) for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

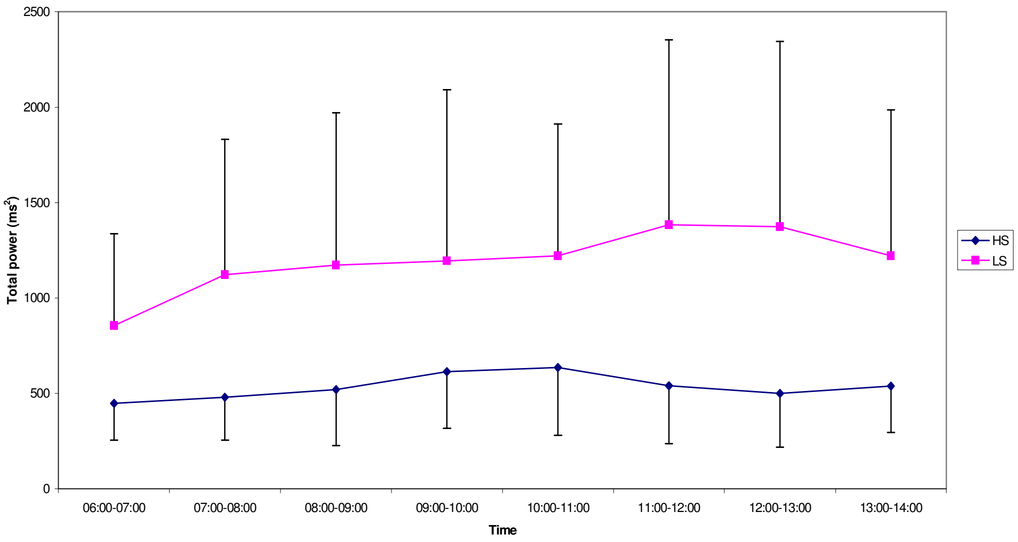


Figure 3.7: Mean total power for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

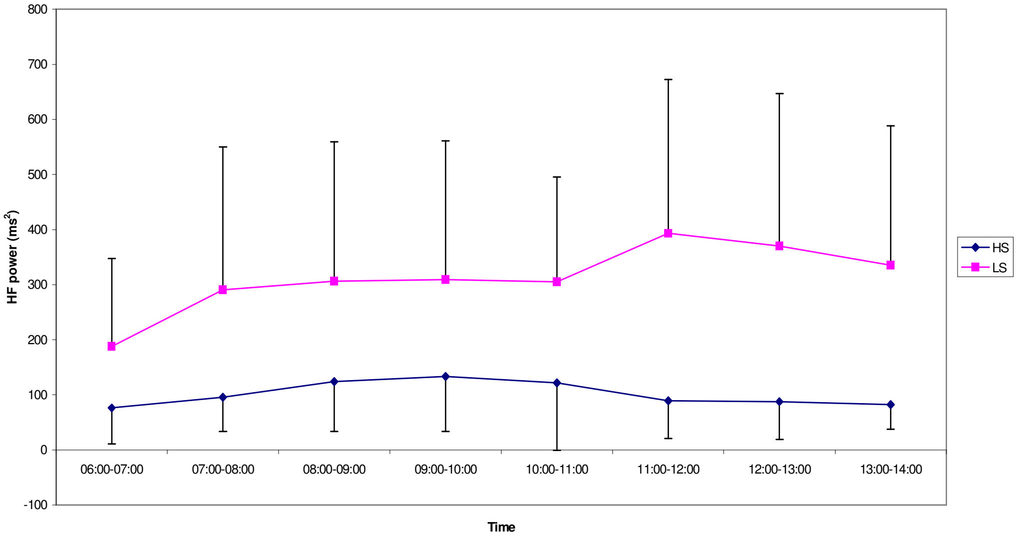


Figure 3.8: Mean high frequency (HF) for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

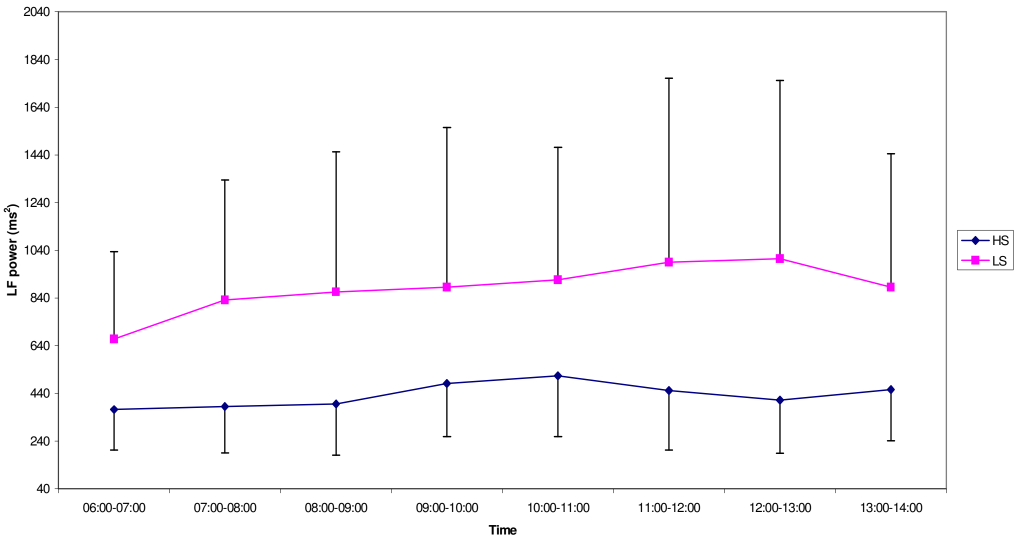


Figure 3.9: Mean low frequency (LF) for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

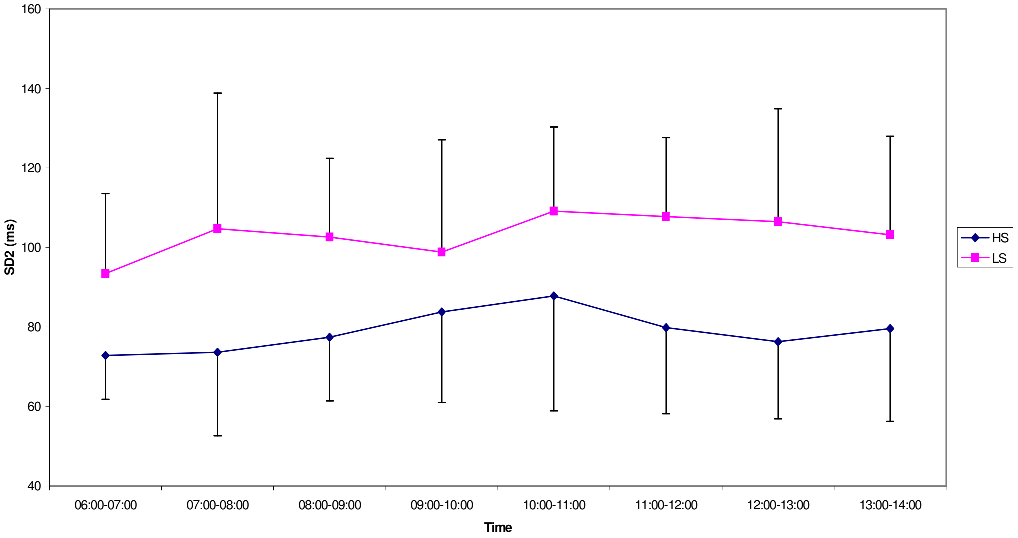


Figure 3.10: Mean SD2 for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

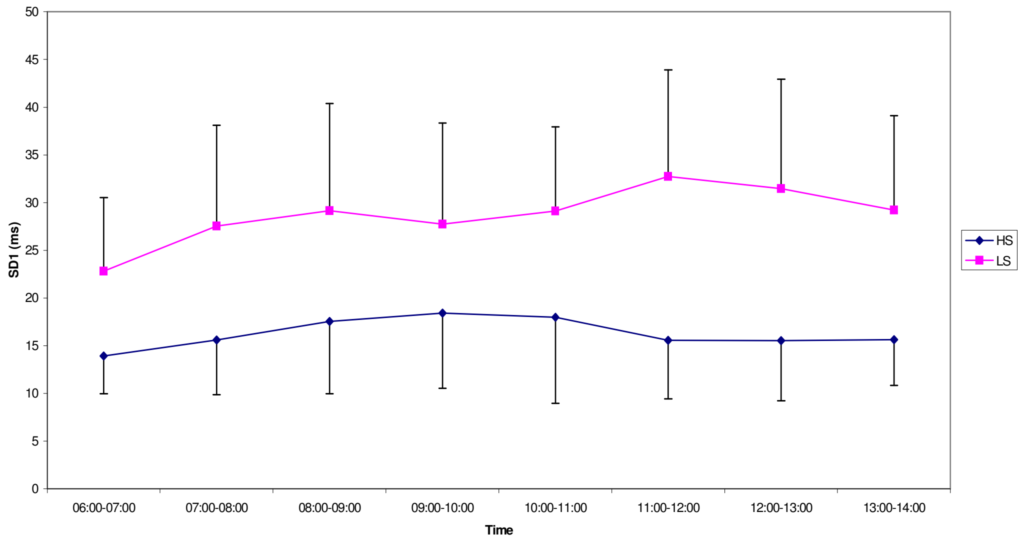


Figure 3.11: Mean SD1 for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

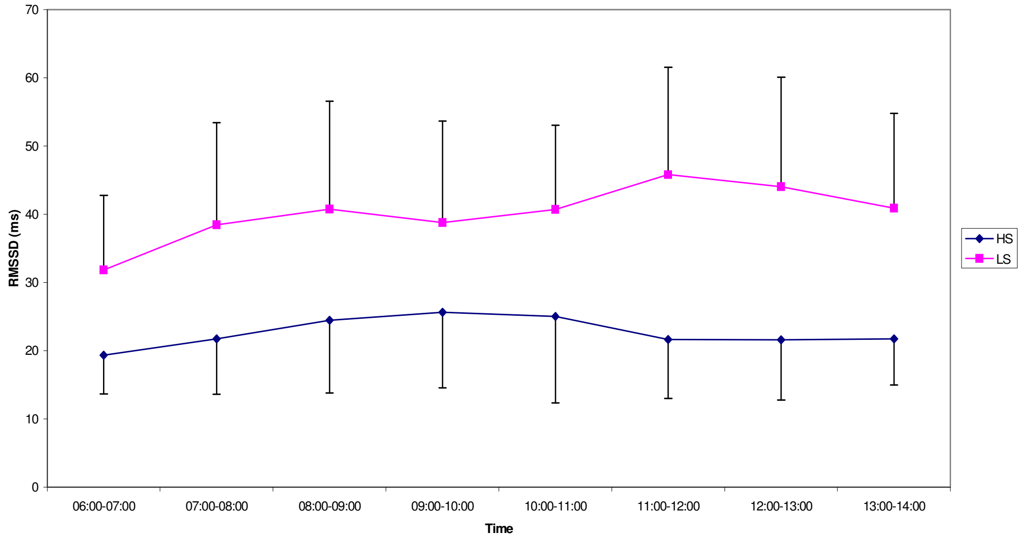


Figure 3.12: Mean RMSSD for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

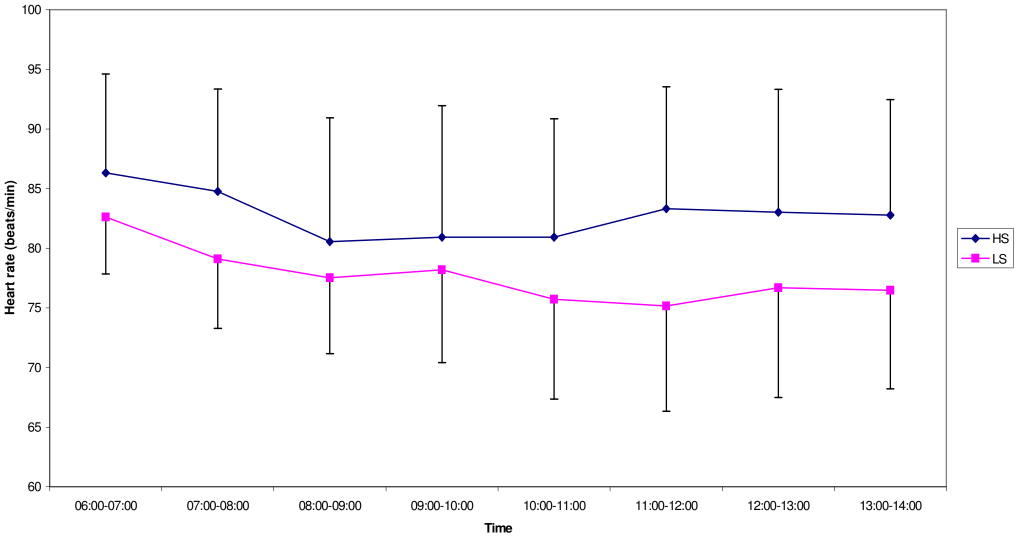


Figure 3.13: Mean heart rate (Minimitter) for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

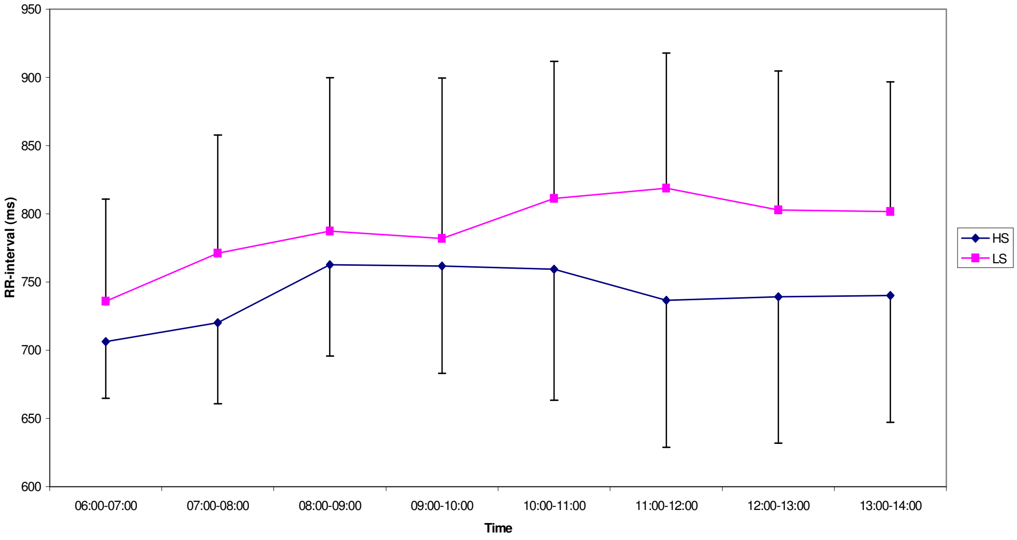


Figure 3.14: Mean RR-interval for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

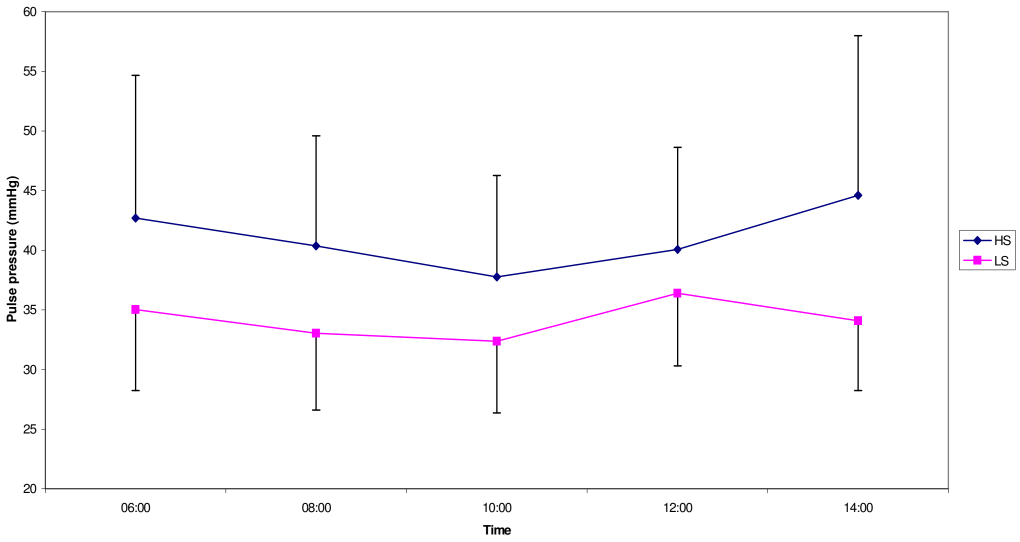


Figure 3.15: Mean pulse pressure for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

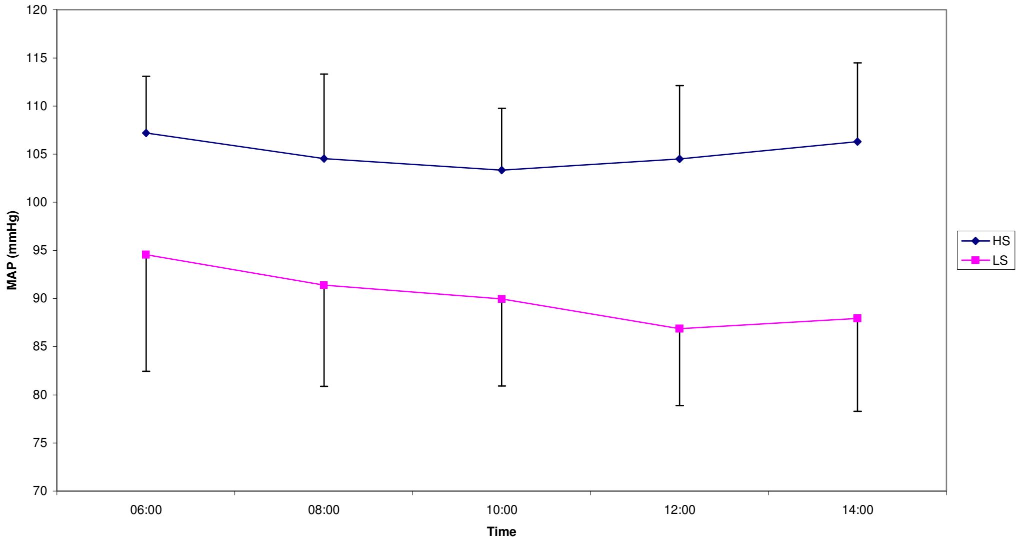


Figure 3.16: Mean arterial pulse pressure (MAP) for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

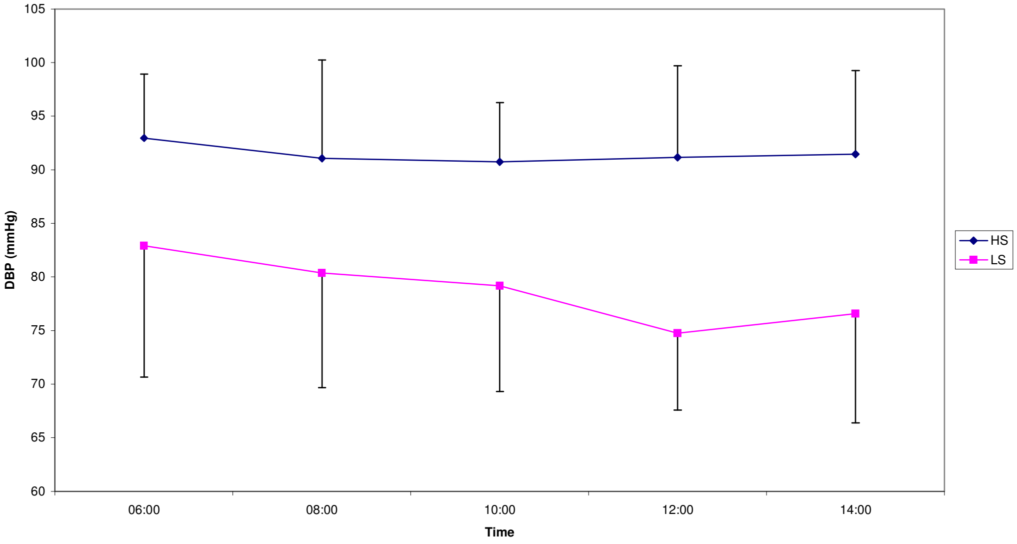


Figure 3.17: Mean diastolic blood pressure (DBP) total number of authorizations for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

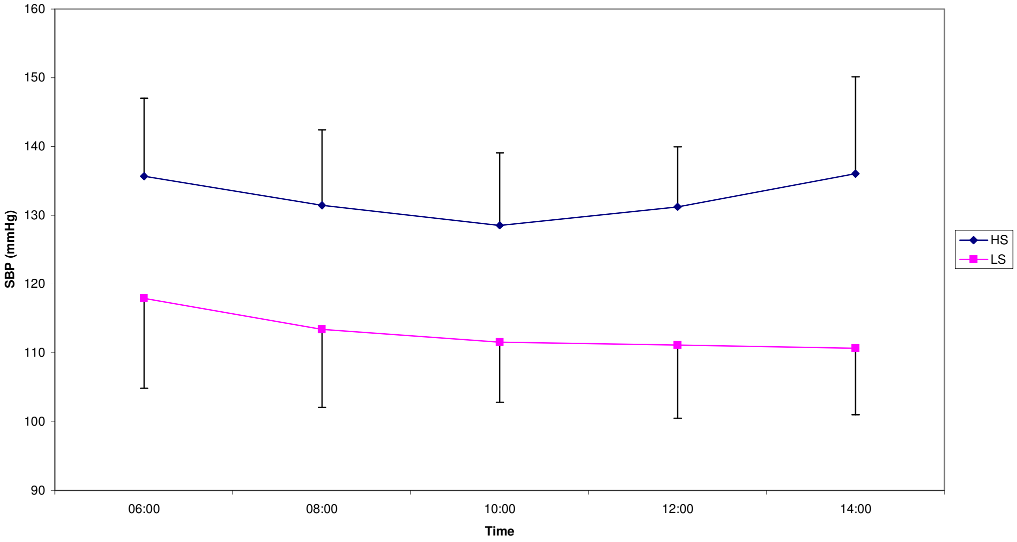


Figure 3.18: Mean systolic blood pressure (SBP) for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

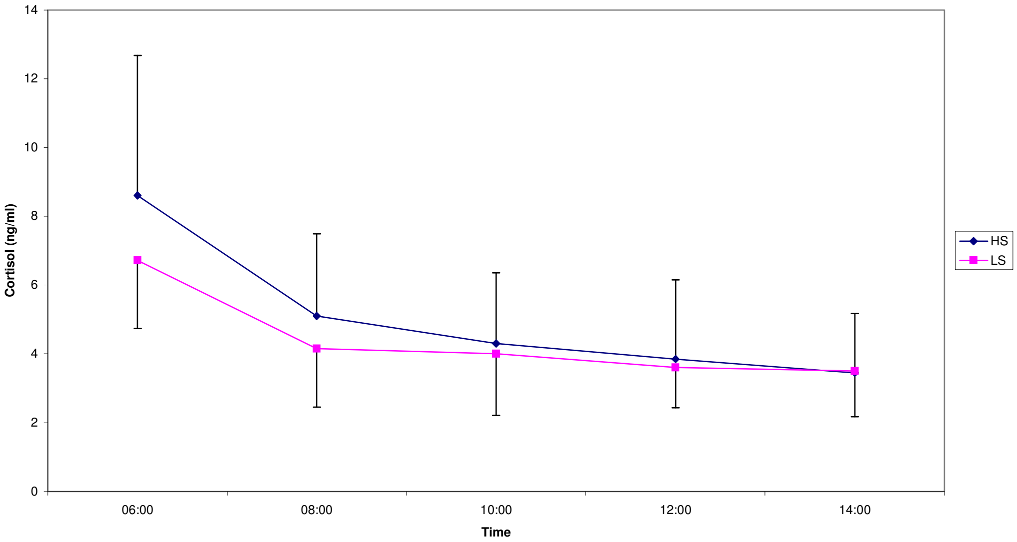


Figure 3.19: Mean cortisol for TCOs designated to fall into the high (HS) or low (LS) stress group according to the Allostatic load index (Model B).

3.4 Integration of MWL-index (Model 2) and Combinations of Experimental Test Parameters and Online Time Analysis

Table 4.1 shows a comparison between the values of the six work stations with the highest and that of the six lowest MWL-indices where the experimental means for the individuals at the six low and six high workstations are calculated form the experimental values over the total work shift.

The evaluation of Models 1 and 2 against combinations of physiological stress parameters (allostatic factors) and against online perception of observer can be seen in Table 2.

Table 4.1: Selection of the six highest and six lowest tables according to the MWL-index (Model 2). Means represent values of subjects grouped into high and low stress groups according to the Physiological Allostatic load index. Values represent means over a six hour period (08:00 – 14:00).

Variable	HS	LS	p value
Age	43.2 ± 5.4	39.1 ± 11.6	0.4893
Mass	106.6 ± 31.5	78.2 ± 19.9	0.0832
Height	1.7 ± 0.1	1.7 ± 0.1	0.8032
Body mass index	35.6 ± 8.3	26.5 ± 3.8	0.0689
Surface area	2.27 ± 0.39	1.93 ± 0.32	0.1317
Cohens	22.6 ± 7.5	16.9 ± 6.7	0.1923
Diastolic blood pressure	95.9 ± 1.7	78.2 ± 6.4	0.0002
Systolic blood pressure	134.6 ± 3.3	112.0 ± 9.6	0.0015
Mean arterial pressure	108.8 ± 1.0	89.4 ± 7.4	0.0004
Pulse pressure	38.7 ± 4.6	33.8 ± 3.8	0.0843
Heart rate (blood pressure monitor)	76.5 ± 3.8	68.6 ± 5.6	0.0357
Cortisol	4.1 ± 1.7	4.0 ± 1.3	0.9270
Heart rate (MiniMitter)	85.3 ± 6.6	80.6 ± 5.0	0.1908
SD of heart rate	4.6 ± 1.1	5.6 ± 1.1	0.1482
RR	714.0 ± 56.5	755.1 ± 49.4	0.2093
SD of RR	33.0 ± 9.4	45.4 ± 10.5	0.0622
RMSSD	21.9 ± 10.6	37.8 ± 10.0	0.0247
SD1	15.7 ± 7.6	27.0 ± 7.1	0.0246
SD2	81.5 ± 26.4	94.6 ± 15.1	0.2987
LF	426.4 ± 188.3	688.9 ± 329.7	0.1427
HF	95.3 ± 82.8	286.6 ± 204.4	0.0544
Ratio (LF/HF)	6.0 ± 2.8	2.7 ± 1.0	0.0597
Total power	521.7 ± 253.7	975.5 ± 510.6	0.0995
Smoking	0.4 ± 0.5	0.6 ± 0.5	0.5995
Length of previous shift	10.4 ± 2.2	10.6 ± 1.8	0.8850
Length of test shift	11.2 ± 1.8	10.6 ± 1.8	0.5648
Years experience at particular station	4.9 ± 4.6	7.4 ± 8.8	0.5840
Shift preference	1.2 ± 0.8	1.0 ± 1.0	0.7234
Radio communications	163.8 ± 100.5	92.3 ± 72.7	0.1811
Telephone communications	50.4 ± 27.2	32.1 ± 20.8	0.2158
Scheduling	58.4 ± 36.4	39.3 ± 33.3	0.3672
Number of Trains	28.0 ± 31.5	15.0 ± 12.9	0.4228
Number of Authorisations	14.8 ± 10.2	16.6 ± 14.3	0.8185

Table 4.2: Evaluation of MWL-index (Model 2) against physiological stress indicators and against online stress perceptions of observer with observer blind to MWL-index classification.

Id	Place	MWL index (Model 2)	Spoornet expert rating	Observer rating	Allostatic index			Agreement				Remarks	Contribution to increased stress	Medication
					BMI, cortisol, sys, dia, hr, lf (A)	BMI, cortisol, sys, dia, hr (B)	cortisol, sys, dia, hr (C)	Observer	Model A	Model B	Model C			
1	[redacted]	Low	Medium	High	Low	Low	Low		1	1	1	Confident, Experience, take no nonsence	Attitude of drivers, radio communication	None
2	[redacted]	Low	Low	Low	Low	Low	Low	1	1	1	1	Nothing to do at all, quiet place	Radio communication	None
3	[redacted]	Low	Low	Low	Low	Low	Low	1	1	1	1	Joked with colleagues, Pleasant atmosphere	Radio communication	None
4	[redacted]	Low	Low	Low	High	High	High	1				Lost temper during shift, highly strung		None
5	[redacted]	High	High	High	High	High	Low	1	1	1		Etopic beats HRV data bad	RIMAS, train problems	None
6	[redacted]	Low	Medium	Low	Low	Low	Low	1	1	1	1	Apprehensive worried about outcome of test		None
7	[redacted]	Low	Low	Low	High	High	High	1				Power failure, Top management NOSA present		None
8	[redacted]	Low	High	Low	High	High	High	1				Cool cat, Open minded / no hang-ups		None
9	[redacted]	High	High	High	High	High	Low	1	1	1		Stop testing at 10 o'clock	Attitude of drivers	None
10	[redacted]	High	Medium	Low	Low	Low	Low						Radio communication	None
11	[redacted]	High	High	High	High	High*	High*	1	1	1	1		Drives all want to talk simultaneously	Blood pressure medication
12	[redacted]	High	Low	High	High	High	High	1	1	1	1	Busy	Drives all want to talk simultaneously, drivers don't wait	Asthma pump
13	[redacted]	Low	Low	Low	High	Low	Low	1		1	1	Chat a lot, no real work relaxed atmosphere	Experience of the people he works with	None
14	[redacted]	High	High	High	High	High	High	1	1	1	1	Fell asleep, relaxed, cool customer	Drivers not co-operative, locomotives to old - when they break the hold up everything	None
15	[redacted]	Low	Low	Low	High	High	Low	1			1	Dubble by pass (1999) high blood pressure medication		Blood pressure medication
16	[redacted]	Low	Low	Low	High	Low	Low	1		1	1	Quiet	Drivers are in a hurry and want authorisations straight away	Blood pressure medication, rheumatism
17	[redacted]	High	Low	High	Low	Low	Low	1				CTC doing TCO work	Broken locomotives, accidents, and radio communication	None
18	[redacted]	High	Low	High	High	High	High	1	1	1	1	Nervous, highly strung, on the edge	Number or trains and work teams, broken locomotives, occupations on line	Blood pressure medication
19	[redacted]	High	High	High	Low	Low	Low	1				Lady, low blood pressure, after 7 very busy, sort out business with section manager	Drivers who don't communicate properly	None
20	[redacted]	High	High	High	Low	Low	Low	1				Lady, low blood pressure, Dolly visited	Radio communication	None
	Low	10		10	8	10	13							
	High	10		10	12	10	7							
	Accuracy (%)							90	50	60	55			

Table 4.3: Pearson correlation of MWL-index (Model 2) with time line analysis variables

Variable	MWL-index (Model 2)	
	r	P
Trains + Authorisations	0.7986	0.0001
Telephone + Radio + Scheduling	0.9110	0.0001
Trains	0.6401	0.0024
Authorisations	0.7286	0.0003
Telephone	0.6306	0.0029
Radio	0.8813	0.0001
Scheduling	0.8342	0.0001

4. DISCUSSION

The brief for this project was to investigate whether the MWL-index, developed by Spoornet, is supported by the values of physiological stress indicators. This was done in terms of changes over the work shifts and in terms of indicators of physiological wear and tear (allostatic load) as a result of long-term exposure to high workloads.

The mental workload index, as developed by Spoornet, consisted of three weighted task factors and eleven weighted moderators. The moderators were factors that carry certain weights (as percentages) as previously decided by a panel of user experts. Tasks factors of the MWL-index were a) the number of data transactions captured by the TCO (weight = 1), b) the number of authorisations (weight = 15) and the number of radio and telephone communications (weight = 5). The moderating factors included a) shift type (shifted work weight = 12%), b) experience as a TCO on the particular system (experience in years on RTO/TWS = 18%), c) planning complexity (interface complexity = 5%, running times between crossing places = 7%, types of crossing places = 5%, location of platforms = 3%, authorisations per shift versus number of crossing places = 10%, d) inherent difficulty (type/mix of trains = 7%, presence of locomotive depots = 9%, presence of shunting yards/activities = 14%, topography = 4%).

In evaluating the MWL-index (Model 1) it became clear that the subdivisions into low and high workload stations, as indicated by this model (Model 1), could not summarily be used as received. The calculations, and therefore the classification of workstations, were based on historical data not reflecting the present workloads at the different venues. A revised classification into work intensity venues was subsequently compiled based on real time data. When comparing the historical classification (based on Model 1), to the real time classification (based on Model 2) there was a difference in the spreading of the various venues over the work intensity spectrum, i.e., the subdivision into high and low centres (Table 2.2). The next step was to compare the subdivisions of workstations as done by the MWL-index to the subdivisions based on stress levels and reactivity in the individuals working at the various stations. In the introduction a background to the rational for measuring the specific stress indicators in individuals was given and it will therefore not be discussed at this point. The mean experimental (physiological) values over the work shifts for sub groupings according to Model 2 were compared (see Table 2.3) to see whether physiological differences could be found between individuals grouped into high and low MWL-indices. The parameters which were included in this comparison comprised factors which could either reflect the stress reactivity in individuals, or the wear and tear as a result of chronic exposure to stressors, or factors that could influence the afore mentioned two types of indicators. Comparisons were made for age, mass, height, body mass index (BMI), surface area (SA), systolic blood pressure, diastolic blood pressure, mean arterial pressure, pulse pressure, heart rate variability variables, smoking or not, length of previous shift, length of test shift, number of years experience a particular station, shift preferences and time line analyses of the shifts when the physiological recordings were made. Very few parameters showed any significant difference between the high and low stress groups. In fact, only the time line analysis, which is anyhow built into the MWL-index, showed, as would be expected, statistical differences between the high and low workload groups throughout (radio communications $p = 0.0003$; telephone communications $p = 0.0276$; scheduling $p = 0.0001$; number of trains $p = 0.0105$; number of authorisations $p = 0.0002$). The influence of the duration of previous shifts differed significantly ($p = 0.0487$) between the two stress groups with longer duration of the previous shifts correlating with the higher stress group. In short, the values of the physiological stress parameters did not mirror the MWL-index model. To see whether exposure to high workloads caused higher expectation stress in workers, the arrival values (before the workload could have had any effect) between the two groups were compared for those parameters measured at arrival. No significant difference could again be found between high and low stress groups (Table 2.4).

In order to sharpen the division between high and low workloads, as calculated by the MWL-index (Model 2), the values of only those individuals at the six highest and six lowest ranked workload stations were compared. These results were seen in Table 2.5. The statistical analyses of this data once again showed that physiological stress values did not mirror the activity intensity as described by the MWL-index. Very significant correlations were again seen between time line analyses and experimental values. These correlations between the time line analysis factors and the early morning value was most probably only a reflection of the fact that time line analyses were built into the MWL-index. It nevertheless confirmed the fact that the MWL-index gives a good reflection of activities at the specific workstations.

The next step was to look at the reactivity over the duration of shifts. When the high and low MWL-index groups, according to MWL-index Model 2, were compared in terms of the way they physiologically reacted to the workload over the total work shifts, hardly any correlations were seen between the workload and reactivity. The results of these repeated measure evaluation between the two groups, change over working time for the whole group, and the group time interaction for the physiological parameters and the time line analysis variables were seen in Table 2.6. Many factors probably contributed to the fact that the results of the MWL-index were not mirrored by the stress levels of the workers. The most likely contributing factors are probably background differences such as age, health status, years of experience and gender. This was discussed under 1.4.4. of the introduction. The populations at the various stations differed significantly with regard to such aspects, and with the experimental group size of this project, such differences could nullify any significant differences. An additional confounding factor is the fact that only three of the twenty workers were female. As discussed in the introduction there are indications that females are perhaps more stress responsive than males. This, however, is contradictory as there are certain factors that protect females against the negative effects of high stress system activation. In this work the females had lower stress levels but were also younger than the average TCO – a fact that could very well underlie their low allostatic loads.

There are several other reasons why workloads of stations would not necessarily be reflected by the physiological values of workers and by changes in response to workload at the specific stations. In the introduction, under adrenal medulla (1.3.3), it was discussed how habituation to known stressors may occur with repeated exposure to the same stressor and how TCOs may therefore adjust to high work loads if subjected to it over extended periods of time. It is only when novel stressors such as dramatic changes to the normal routine, are introduced into the work environment that sensitisation and the effect of high allostatic loads with subsequent influences on their skills, as well as their mental and physical health might be noticed. It speaks for itself that changes in infrastructure and policies may resort under such novel environmental stressors. A similar phenomenon was also discussed in the introduction under HPA-axis (1.3.4 and 1.4.3). To quote "although acute stress can cause an increase in the basal-to-peak difference of cortisol the amplitude of the rhythm decreases in conditions of chronic stress as a result of the increase in baseline levels. As in the case of the SAM-axis, bouts of heterotypic acute stressor application during periods of chronic stress alter the response. It would seem that acute stress superimposed on chronic stress is dependent on the familiarity with the type of stressor. As in the case of the SAM-axis, it appears that the response to an acute stressor in the chronically stressed would be less than expected if the stressor is homotypic (a repetition of what caused the chronic stress response). In contrast, if a heterotypic stressor is applied to a chronically stressed individual, the response would be bigger. It is however possible that cross-tolerance to stressors may develop – especially stressors that involve the same neurological pathways. This once again can in theory be extrapolated to the working situation where high cortisol levels would then not be as significantly increased by increased levels of homeotypic stressors such as increases in the amount of work to which the individual is accustomed. In contrast, the response can be exacerbated when stressors other than the typical work stressors to which the individual is accustomed to are encountered. The implication that can be deduced from the previously mentioned fact (i.e., that the amplitude of the stressor-induced increase in cortisol levels, as well as the circadian amplitude difference to additional homeotypic stressors may be lower during chronic high stress) would be that those individuals with chronic high cortisol levels will not show the expected increase from basal trough values when stress levels increase and that this could be ascribed to the fact that the trough values are higher than normal". To extrapolate this to the TCOs, one could expect that habituation to their respective workloads could rule out the development of significant stress differences between the different work intensity groups and that the magnitude of the stress reactions from baselines to stimulated responses could be negated by above normal baseline values. The latter part of the last statement would apply to stressor-induced changes over shifts.

The next step was to identify high and low stress groups on the grounds of physiological variables. After the comparison between the values of the Spoornet index of workloads at the various stations on the one hand, and the physiological indicators of stress in the workers at the different stations/tables on the other, the possibility of separating the workers at the different stations into high and low stress groups on grounds of the values of their physiological parameters was investigated. The analyses, as previously mentioned, were performed in terms of an adapted allostatic load measurement where the data used in the calculation of the numerical value for each individual factor or parameter of the allostatic load were derived from the mean of the values obtained over the work shift. The reason for this was that baseline values in this group could not really be seen as baseline as all of them appeared to arrive with values higher than any reached during the experimental procedure. Many factors could have contributed to this type of observation, not least of all the fact that the arrival time coincided with a steeper part of the circadian rhythm of factors such as cortisol, while subsequent measurements were performed at times when the circadian values were already lower. Expectation or anticipation of the work stress before initiation of the work shift could further have contributed. Anticipation as a stressor and the orienting reflex were discussed in the introduction under 1.6.2 and 1.6.3. In addition, it is known that individuals in high stress jobs seldom recover to baseline value over the time off. This has been shown by various studies such as that of Steptoe, et al 1999, (32). In the study on job strain of Steptoe et al, blood pressure, heart rate and electrodermal responses were determined to externally paced (uncontrollable) and self-paced (controllable) tasks. The results of the above study showed that blood pressure reactions to uncontrollable tasks were greater in high than low job-strain groups, but the same differences were not seen with controllable tasks. Systolic and diastolic pressures did not differ between their groups over the day but decreased more over the evening with low job strain. The authors concluded, and rightfully so, that the failure of high strain job subjects to reduce their blood pressure over the evening may be a manifestation of high allostatic loads. There can be no doubt that the stressors in the case of TCOs should be considered uncontrollable as the number of trains and other activities required are predetermined by the job at hand and is not under their own control. The fact that the stressor, in this case, is not a once off stressor but that TCOs exist with these levels of stress axes activity for the major part of every day supports the decision to use the mean value over the work shift.

As previously mentioned, in the process of subdividing workers at various stations/tables into high and low stress groups, based on physiological values, the values of the following parameters were included: salivary cortisol and BMI as indices of hypothalamo-pituitary-adrenocortical axis activity, systolic and diastolic pressure as measure, of cardiovascular activity – largely reflecting sympatho-adrenomedullary axis activation, heart rate as an indicator of sympatho-adrenomedullary activation and heart rate variability variables as indicators of autonomic activity. The values of each individual for each of the 5 indicators were classified according to the 50th percentile. Allostatic load was then calculated by summing up the number of parameters for which the subjects fell into the highest risk, i.e., above the 50 percentile. Subjects were subsequently ranked according to their total hits, as previously discussed. Three models, Model A (including the values of cortisol, systolic blood pressure, diastolic blood pressure, heart rate, low frequency and BMI), Model B (cortisol, systolic, diastolic, heart rate and BMI) and Model C (cortisol systolic, diastolic and heart rate) were developed and tested and the TCOs subdivided into high or low stress groups. When the TCOs were subdivided into high and low stress groups and the differences for all physiological parameters (not only those included in the three models, respectively were tested for significant differences between the groups of individuals subdivided into high and into low according to Model A, B and C, it was seen that Model B was superior to Model A and C. This was supported by the fact that for Model A 10 out of 17, for Model C 8 out of 17, and for Model B 12 out of 17, physiological variables differed significantly ($p < 0.05$) between the high stress groups and low stress groups (two sample T-test). The significances were further generally higher for Model B than for the other 2. Model B also divided the 20 TCOs directly into a 50-50 split. The superiority of model B was also borne out by the results of the Two-way AOV where heart rate variability variables confirmed significant differences between the two stress groups in terms of statistical, frequency domain and geometric methods of analyses (total power, low frequency power, root means square of successive differences, standard deviation of RR, standard deviation of HR, etc).

In the final analysis, subdivisions of workstations on the basis of the MWL-index, were compared to subdivisions of the individuals at those stations, and to online perceptions of the observer about the stressor value of the workstation. A 90% agreement was seen between the MWL-index and online observer perception (this was subjective perception and not based on the sum of the activity as determined by time line analyses) and a 60% agreement between the subdivisions into high and low stations by MWL-index on the one hand, and subdivisions based on Model B on the other.

Time line analysis is built into the MWL-index and one would therefore expect to find some kind of correlation between time line analysis and MWL-index. In an attempt to investigate the strength of the correlation and to see whether the MWL-index has a significant advantage over simple time line analysis, correlations was tested between the MWL-index on the one hand and the individual factors in time line analysis, as well as two combinations of time line analysis factors (Table 4.3). The correlations between the MWL-index and the two combinations of factors were $r = 0.7986$; $p = 0.0001$ (trains and authorizations) and $r = 0.9110$; $p = 0.0001$ (radio, telephone, schedules) and all individuals factors also showed significant correlations. It would therefore appear that the time line analysis very much gives the same information as that derived from calculating the model.

5. SUMMARY

The original subdivisions of stations into high and low workload, was based on historical data (Model 1). A revised MWL index (Model 2) was created to incorporate real time data (time line analysis) recorded during physiological testing. In using the revised MWL-index it was shown that it is imperative to incorporate results of the latest time line analysis, and not historical data, in the estimation of the workload at the various stations. The workloads, at the various stations, as predicted by the MWL-index (Model 2), were not reflected by either the adapted allostatic load (including all measured parameters of the individuals working at those stations), or by changes in stress levels over work shifts. However, in developing three models, consisting of different combinations of allostatic load indicators, there was a 60% correspondence between subdividing workstations into low and high workloads according to the MWL-index, and subdivision of workers at the corresponding stations into high and low stress according to Model B. Model B was based on the means of the values taken over the shift for cortisol, systolic pressure, diastolic pressure, heart rate and BMI. It can thus be said that the combination of physiological parameters used in Model B supports the validity of the MWL-index as indicator of work stress at the various stations. In the final analysis it is necessary to ask whether either the development of a Mental Workload Index or the use of physiological parameters of the workers at the various stations gave significantly better estimates of the stress levels at the stations than the mere use of simple time line analysis. At this stage the answer will have to be in the negative. Although the MWL-index is a good reflection of the activities at the stations, and Model B supports the use of the index, neither would appear to have an advantage over simple time line analysis over the work shift. Another question, in view of the high correspondence between the MWL-index and subjective perception by an observer, is whether the workload cannot simply be estimated through observance without either time line analysis, physiological measurements or calculation of the MWL-index. Once again the answer should be in the negative as it could, depending on the observer, change with the ability and commitment of the observer as well as her or his objectivity.

6. CONCLUSIONS AND FINAL REMARKS

- The MWL-index as received from Spoornet (Model 1) used historical data (2003) instead of real time data. The model was corrected for real time data and the corrected version (Model 2) thereafter used as the MWL-index.
- The total spectrum of physiological stress indicators did not support the workload as predicted by the corrected MWL-index. However a combination of physiological stress indicators (referred to as Model B) did support (60% correspondence) the classification into high and low workstations as predicted by the MWL-index.
- Although the MWL index (Model 2) gave a relatively good reflection of the activities at the stations, and the allostatic load index (Model B) supports the use of the MWL index, neither appears to have an advantage over simple time line analysis over the shift.
- Factors such as habituation and sensitisation influence physiological responses to work stress and could have attenuated the correlations between the workloads and the physiological responses. Habituation (decrease physiological responses to work stress) may be disturbed by marked changes in the work environment or company policies and may once again lead to overt stress responses in the face of high workloads.

7. REFERENCES

1. Stress, cognition and health. Psychological focus. Tony Cassidy. Routledge, London and New York. 1999, pp1-14.

2. Stress, coping and development. An integrative response. Caroline Aldwin. The Guilford Press, New York, London, 1994.

3. Viru A. Mechanisms of general adaptation. Medical Hypothesis 1992;38:296-300.

4. Chrousos GP, Gold PW. The concepts of stress and stress system disorders. JAMA 1992;267:1244-1252.

5. Lovallo WR. Stress & Health: Biological and Psychological Interactions. SAGE Publications, California, 1997.

6. Kvetnansky R. Adrenal medulla. In: George Fink (Editor-in Chief), Encyclopedia of stress Volume 1, Academic Press, London. 2000, pp63- 70. In: George Fink (Editor-in Chief), Encyclopedia of stress Volume 2, Academic Press, London. 2000, pp224-237.

7. Dallman MF, Bhatnagar S, Viau V. Hypothalamo-pituitary-adrenal axis. In: George Fink (Editor-in Chief), Encyclopedia of stress Volume 2, Academic Press, London. 2000, pp468-483.

8. Dallman MF. Glucocorticoid negative feedback. In: George Fink (Editor-in Chief), Encyclopedia of stress Volume 2, Academic Press, London. 2000, pp224-228.

9. Buckingham JC. Effects of stress on glucocorticoids. In: George Fink (Editor-in Chief), Encyclopedia of stress Volume 2, Academic Press, London. 2000, pp229-237.

10. Bernson GG, Cacioppo JT. From homeostasis to allodynamic regulation. In: John T Cacioppo, Louis G Tassinary, Gary G Berntson. Handbook of psychophysiology (2nd edition), Cambridge University Press, Cambridge 2000, pp459-481.

11. McEwen BS. Allostasis and allostatic load. In: George Fink (Editor-in Chief), Encyclopedia of stress Volume 1, Academic Press, London. 2000, pp145-150.

12. McEwen BS. Seminars in medicine of the Beth Israel Deaconess Medical Center: protective and damaging effects of stress mediators. The New England Journal of Medicine 1998;338(3).:171-179

13. Sterling P, Eyer J. Allostasis: A new paradigm to explain arousal pathology. I: S Fisher and J Reason (eds). Handbook of life stress, cognition and health, Wiley, New York, 1988 pp629-649.

14. Seeman TE, Robbins RJ. Aging and hypothalamic-adrenal responses to challenge in humans. Endr Rev 1994;15:233-260.

15. Albeck DS, Mc Kittrick CR, Blanchard DC, Nikolina J, McEwen BS, Sakai RR. J Neurosci 1997;17:4895-4903.

16. Jefferies WM, Mild adrenocortical deficiency, chronic allergies, autoimmune disorders and the chronic fatigue syndrome: a continuation of the cortisone story. Med Hypoth 1994;42:183-189

17. McEwen BS. Allostasis and allostatic load. In: George Fink (Editor-in Chief), Encyclopedia of stress Volume 2, Academic Press, London. 2000, pp145-150.

18. Jennings JR. Heart rate. In: George Fink (Editor-in Chief), Encyclopedia of stress Volume 2, Academic Press, London. 2000, pp333-337.

19. 18 Woo MA, Stevenson WG, Moser DK, Trelease RB, Harper RM. Paterns of beat-to-beat heart rate variability in advanced heart failure. American heart journal 1992;March:704-710.

20. Human Physiology (3rd Edition): BJ Meyer, DH van Papendorp, HS Meij, M Viljoen. Juta, Pretoria, 2002, pp14.1-14.11

21. Sherwood A, Carels RA. Blood pressure. In: George Fink (Editor-in Chief), Encyclopedia of stress Volume 1, Academic Press, London. 2000, pp331-338.

22. Manuck S B, Kasprowicz AI, Muldoon MF. Behaviouraly evoked cardiovascular reactivity and hypertension: Conceptual issues and potential associations Ann Behav Med 1990;12:17-29.

23. Brownley KA, Hurwitz BE, Schneiderman N. Cardiovascular psychophysiology. In: John T Cacioppo, Louis G Tassinary, Gary G Berntson. Handbook of psychophysiology (2nd edition), Cambridge University Press, Cambridge 2000, pp224-264.

24. West S, Stanwyk C, Brownley K, Bragdon E, Hinderliter A, Light K. salt restriction increases norepinephrine and blacks more than in whites [abstract]. Ann Behav Med. 1997;19:S067.

25. Ironson GH, Gellman MD, Spitzer SB, Llabre MM. Pasin RD, Weidler DJ, Schneiderman N. predicting home and work blood pressure measurements for resting baseline and laboratory reactivity in black and white Americans. Psychophysiol 1989;26:174-184.

26. Lovallo WR, Thomas TL. Stress hormones in psychophysiological research. Emotional, behavioural, and cognitive implications. In: John T Cacioppo, Louis G Tassinary, Gary G Berntson. Handbook of psychophysiology (2nd edition), Cambridge University Press, Cambridge 2000, pp342-367.

27. Schneiderman N, McCabe PM. Psychophysiologic strategies in laboratory research. In: N Schneiderman, SM Weiss, PG Kaufman (eds), Handbook of research methods in cardiovascular behavioural medicine, 1989, Plenum, New York, pp349-364.

28. Schulkin J, McEwen BS, Gold PW. Allostasis, amygdala and anticipatory angst. Neurosci & Biobehav Rev 1994;18(3):385-396.

29. Pavlov IP. Conditioned reflexes. Oxford University Press, Oxford, 1927.

30. Sokolov EN, The neural mechanisms of the orienting reflex. In: EN Sokolov, OS Vinogradova (eds), neuronal mechanisms of the orienting reflex. NJ: Erlbaum, Hillsdale,1975, pp217-235.

31. Wallin BG. Sympathetic nerve activity underlying electrodermal and cardiovascular reactions in man. Psychophysiology 1981;18:470-476.

32. Steptoe A, Cropley M, Joekes K. Job strain, blood pressure and response to uncontrollable stress. Journal of Hypertension 1999;17(2):193-200.