U.S. Department
of Transportation

Federal Railroad Administration

A Preliminary Examination of Railroad Dispatcher Workload, Stress, and Fatigue

Office of Research
and Development
Washington, DC 20590

S. Popkin
J. Gertler
S. Reinach

Foster-Miller, Inc.
350 Second Avenue
Waltham, MA 02451-1196

DOT/FRA/ORD-01-08

Final Report
May 2001

This document is available to the U.S. public through the National Technical Information Service, Springfield, VA 22161

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle A Preliminary Examination of Railroad Dispatcher Workload, Stress, and Fatigue		5. Report Date May 2001 6. Performing Organization Code
7. Author(s) S. Popkin, J. Gertler, S. Reinach		8. Performing Organization Report No. DOT/FRA/ORD-01-08
9. Performing Organization Name and Address Foster-Miller, Inc. 350 Second Avenue Waltham, MA 02451-1196		10. Work Unit No. (TRAIS) 11. Contract or Grant No. DTFR-53-95-00049
12. Sponsoring Agency Name and Address U.S. Department of Transportation Federal Railroad Administration Office of Research and Development Washington, DC 20590		13. Type of Report and Period Covered March 1996 – September 1998 Final Report 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

This report presents the methods, findings and recommendations from a field study that examined the sources and levels of railroad dispatcher workload, stress, and fatigue. The study was initiated in response to concerns raised by two Federal Railroad Administration safety audits of dispatching operations in the U.S. The work was performed in three phases. The first phase identified suitable data collection instruments for measuring the three factors, Phase 2 involved a pilot field test of the instruments and field study procedures, and Phase 3 involved field data collection.

Data collected in the field at two sites included physiological measures as well as self-report data, third party observation and paper records. Based on self-report health data, study participants aged 25 to 44 experienced several disorders at a significantly higher rate than found among the U.S. population. Information from the participant background survey was used in conjunction with physiological measures, activity count data and subjective ratings of workload, stress and fatigue. Subjective workload ratings were moderately associated with reported number of trains dispatched, regardless of shift or location while subjective fatigue ratings had the lowest association with this measure of workload. Subjective ratings of stress increased throughout all shifts while salivary cortisol levels, a physiological measure of stress, were within normal adult limits. Sleep patterns of study participants were typical of shift workers in general. An observational technique provided additional data on variations in workload throughout the shift.

Modifications and enhancements to both the measures and protocol for future studies of railroad dispatcher workload, stress, and fatigue are offered as issues for further research. Appendices contain copies of the data collection tools used in the field study along with supporting data and statistical analyses.

17. Key Words	18. Distribution Statement
railroad dispatcher, workload, stress, fatigue	Document is available to the U.S. public through the National Technical Information Service, Springfield, VA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	203	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

PREFACE

This report presents the background, methods, and results of a research study designed to examine the workload, stress and fatigue of railroad dispatchers. The work was performed under contract DTFR53-95-C-00049 for the Federal Railroad Administration. Dr. Thomas Raslear, Office of Research and Development, Federal Railroad Administration, was the contracting officer's technical representative (COTR). The authors wish to thank Dr. Raslear for his valuable guidance throughout the conduct of the study. The authors extend special thanks to Mr. Thomas Keane, Office of Safety, Federal Railroad Administration, who provided a background on the FRA's two dispatcher audits and answered numerous questions throughout the course of the project.

Thanks are also due to the management of the two railroads that allowed data collection to take place in their facilities, the union representatives who facilitated the logistics for field data collection, Mr. William Clifford, former President, Brotherhood of Locomotive Engineers, American Train Dispatchers Department (BLE/ATDD) for his support during the conduct of the study and all of the railroad dispatchers who participated in the study.

In addition to the authors, several other individuals contributed to this research project. Dr. George Kuehn, IIT Research Institute, adapted the Task Analysis Workload (TAWL) methodology to this study, oversaw field data collection and assembled the results. Dr. Theodore Baker, Shiftwork Resources, was responsible for collection of the saliva samples and analyzing them for cortisol levels. Ms. Susan Madigan, Transit Safety Management, provided expertise regarding dispatching operations. Mr. Jeffrey Crane, Foster-Miller, performed dispatcher observation and assisted in coding of the field data. Dr. William Stankard, Mr. Jeremy Mirsky and Ms. Denise Rose also performed dispatcher observation. Ms. Sally Pham, Foster-Miller, assisted with data analysis and preparation of this report.

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in²	square inches	645.2	millimeters squared	mm²
ft²	square feet	0.093	meters squared	m²
yd²	square yards	0.836	meters squared	m²
ac	acres	0.405	hectares	ha
mi²	square miles	2.59	kilometers squared	km²
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	ml
gal	gallons	3.785	liters	l
ft³	cubic feet	0.028	meters cubed	m³
yd³	cubic yards	0.765	meters cubed	m³

NOTE: Volumes greater than 1000 l shall be shown in m³.

Symbol	When You Know	Multiply By	To Find	Symbol
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams	Mg
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5(F-32)/9 or (F-32)/1.8	Celcius temperature	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	l
fl	foot-Lamberts	3.426	candela/m²	cd/m²
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
psi	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm²	millimeters squared	0.0016	square inches	in²
m²	meters squared	10.764	square feet	ft²
m²	meters squared	1.195	square yards	ac
ha	hectares	2.47	acres	mi²
km²	kilometers squared	0.386	square miles	
		VOLUME		
ml	milliliters	0.034	fluid ounces	fl oz
l	liters	0.264	gallons	gal
m³	meters cubed	35.71	cubic feet	ft³
m³	meters cubed	1.307	cubic yards	yd³
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg	megagrams	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact)		
°C	Celcius temperature	1.8C + 32	Fahrenheit temperature	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m²	candela/m²	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	psi

* SI is the symbol for the International System of Units

536-Conversion Factors

(Revised January 1992)

CONTENTS

Section **Page**

Section		Page

Section		Page

ILLUSTRATIONS

Figure **Page**

Figure **Page**

TABLES

Table **Page**

Table **Page**

LIST OF ABBREVIATIONS

ACTH	adreno-corticotrophic hormone
ANOVA	analysis of variance
ATC	air traffic control
ATDD	American Train Dispatcher's Department of the Brotherhood of Locomotive Engineers
CADS	computer-aided dispatching system
CRH	corticotropin releasing hormone
CTC	centralized train control
EKG	electrocardiogram
EMG	electromyogram
GPS	global positioning system
GSR	galvanic skin response
HPA	hypothalmic-pituatary-adrenal
HSD	honestly significant difference
IRB	institutional review board
MCH	Modified Cooper-Harper
MOW	maintenance of way
mTAWL	Modified Task Analysis Workload
NASA TLX	NASA Task Load Index
NIMS	NAS Infrastructure Management System
NPRU	Naval Psychiatric Research Unit
OW	Overall Workload
PTC	positive train control
REM	rapid eye movement
SWAT	Subjective Workload Assessment Technique
T&E	train and engine
TAWL	Task Analysis Workload
VTS	Vessel Traffic Services

EXECUTIVE SUMMARY

Railroad dispatchers are responsible for the safe and efficient movement of trains and other track users over a railroad. In two Federal Railroad Administration safety audits conducted in the 1990s that covered 125 dispatching offices and over 1000 dispatchers, the FRA Office of Safety found evidence of high dispatcher stress and workload. The FRA recognizes the significant role that the dispatcher plays in the safety of railroad operations, and is aware of the significant impact that dispatcher workload, stress and fatigue can have on employee health and well-being, which in turn may create safety risks.

Consequently, the FRA initiated this research to understand today's dispatching environment and its associated levels of workload, occupational stress and resulting fatigue. The goals of this research were to: 1) identify the sources and magnitude of workload, stress and fatigue associated with the dispatcher's job and working life, 2) determine any related health or performance effects, and 3) refine procedures for measuring workload, stress and fatigue in the dispatcher's workplace.

The study had three phases. In Phase 1, appropriate data collection measures were selected to quantify dispatcher workload, stress and fatigue. A pilot field study was conducted in Phase 2 to evaluate the candidate measures and data collection procedures. In Phase 3, based on the results of Phase 2, a full field study, which involved two weeks of data collection, was conducted at two dispatching operations.

A literature review of related studies in other fields, site visits to dispatching centers and a focus group interview with railroad dispatchers all helped to identify and evaluate candidate data collection instruments in Phase 1. Because this study was conducted in the dispatchers' workplace, dispatcher acceptance of the instruments became the key criterion for instrument selection. More specifically, ease of use, time to administer and lack of interference with job duties were considered. Phase 2 of the project involved a one-week pilot study to evaluate selected data collection instruments and procedures. Several minor procedural changes were made based on the pilot study.

The final set of data collection instruments is summarized in Table 1. A background survey collected data on the demographics and health of the study population along with information about the work environment. There were three sources of workload data, including an observational technique, based on the Task Analysis Workload method, subjective ratings and activity count data. Subjective ratings were also used to collect data on dispatcher stress and fatigue. Salivary cortisol was used as a physiological measure of stress while actigraphy was used to record sleep patterns. Dispatchers also provided a record of their sleep pattern on a sleep

Table 1. Data collection instruments

Type of Data	Measurement Instruments
Demographic and Work Environment	Background survey
Workload	mTAWL Activity records Dispatcher self-reports Subjective ratings
Stress	Salivary cortisol Subjective rating
Fatigue	Actigraphy Sleep logs Subjective rating
Study Feedback	Debriefing survey
Job Improvement Suggestions	Debriefing survey

log. Finally, a debriefing survey solicited feedback from participants regarding study procedures and offered the participants an opportunity to express their suggestions for job improvements.

A total of 20 dispatchers from a freight dispatching operation and 17 from a passenger operation participated in the Phase 3 field study. Participants volunteered and were compensated for their participation. Prior to the start of data collection each participant signed an informed consent form and completed the background survey. During the two-week data collection period, participants wore an actigraph and maintained a sleep log on both work and nonwork days. While at work, participants completed subjective rating forms and provided a saliva sample upon arrival at work and every 2 hr thereafter. At the conclusion of the two-week period, each participant completed the debriefing survey and was paid for his or her participation.

No significant differences were found between the two field study locations with respect to dispatcher characteristics. Study participants, predominantly white married men, had a median age of 43 and median job tenure of 8 years. The characteristics of this group likely reflect the largely homogeneous railroad work force in general. With respect to health, the younger dispatchers, aged 25 to 44, reported experiencing back pain, headaches and skin disorders at a significantly higher rate than found in the U.S. population in general. Since these health problems can result from chronic stress, health issues of the entire U.S. dispatcher population may merit further investigation.

The study results indicate that participating dispatchers worked more than a 40 hr week. Most dispatchers (89 percent) reported that, on average, they are scheduled to work a consecutive five-day workweek, but over half reported working an average of eight or more hours of overtime weekly. The dispatcher’s time at work is further stressed by the fact that the dispatcher’s work schedule does not provide a scheduled lunch or rest break. Although participants reported taking a median of four breaks per shift, their responses on the background survey indicated that they are not always taken when desired. The pressure to work overtime reduces the dispatcher’s free time, a situation that is potentially stressful.

Regarding the job and work environment, though nearly all the dispatchers reported that they often dealt with high workloads, most felt capable of handling emergencies and competent in all aspects of their job. Separately, as reported in the background survey responses, stress was attributed less to the work demands and more to issues of control over work and interpersonal interactions.

The modified Task Analysis Workload (mTAWL) measurements provide a means to gauge the variation in workload over the course of a shift. Due to the mTAWL's labor-intensive nature, it may not be suitable for a research study. However, it has potential as a tool for identifying and documenting workload at a specific desk and for comparing workload over time or across desks within a dispatching center. For example, using charts of the mTAWL data, differences in the workload pattern between a Tuesday and a Friday at the freight operation were observed. The surge in workload came earlier on Friday and overall was at a higher level than on Tuesday. Between desk comparisons identified one desk at each operation where the dispatcher was overloaded a high percentage of time relative to other desks at the same operation.

Comparisons of the available activity count data – number of trains and other track users, numbers of Form Ds[1] and route blocks – with subjective ratings of workload, stress and fatigue revealed that subjective workload was moderately associated with reported number of trains dispatched, regardless of shift or location. Correlations between the count data and the perceived workload were significant for 92 percent of the cases. Perceived stress also related to the number of trains dispatched, particularly at the freight operation where 92 percent of the coefficients were significant and reliable compared with 42 percent at the passenger site. Of the three subjective ratings, fatigue had the lowest association with the number of trains and other track users, particularly among the dispatchers at the passenger operation. This suggests that feelings of fatigue are not merely a function of workload as defined by number of trains and other track users handled. Multiple regression analysis to examine the relationship between the subjective ratings and number of Form Ds revealed that while there is a reliable linear relationship between number of Form Ds and route blocks handled and the subjective ratings, the actual impact of the number of Form Ds on the ratings is extremely limited. (The R^2 for these regressions ranged from 0.00 to 0.12.)

The background survey provided information on the dispatchers' general level of stress and sources of this stress. Two distinct patterns of stressors emerged for each site. Dispatchers at the freight operation appeared to be primarily concerned with their workload. Specific workload-related stressors that concerned them included difficulty of work, surges in workload, and lack of control. In contrast, the passenger operation dispatchers found personal interactions and the physical work environment to be their primary stressors. When these responses were analyzed by job tenure, those with two to five years of experience reported the greatest number of stressors. This is likely explained by the fact that this group is expected to perform at the same level as the more experienced dispatchers while they are still building the mental models, heuristics and information that experienced dispatchers already possess.

[1]A Form D is a written or electronic record of track usage authority issued by a railroad dispatcher operating under NORAC rules.

The passenger operation dispatchers were given the opportunity at the end of the study to "List your top five contributors to work stress, most stressful first." Surprisingly only one of the items reported was directly workload-related, number of concurrent tasks. Thus, according to self-report data, stress at these sites appears to be multivariate in nature, and is not completely centered around the work itself.

In contrast to the results of the background survey, there was little evidence of a high level of stress from either the subjective stress ratings or the salivary cortisol levels, the physiological measure of stress. The salivary cortisol results for this group of dispatchers were well within normal levels for adults. The diurnal pattern of their cortisol also matched available adult norms. These results, however, should not be interpreted as an indication that workplace stress does not exist. More likely, data was collected too infrequently in this rapidly changing environment and may not have captured the changing workload and related stress.

Subjective ratings of fatigue significantly increased throughout the duration of all shifts, doubling or nearly doubling in all cases. Fatigue ratings for the start and end of the night shift were significantly higher than those for the day and evening shifts. This is likely due to the dispatchers fighting their circadian rhythm as well as handling the responsibilities related to being on duty.

Sleep patterns of the participants appeared normal and there was no evidence of an acute sleep debt. The rate of those "waking up tired" on workdays and nonworkdays shows a pattern consistent with the shiftwork literature. Specifically, those working nights have the highest rate of this symptom, and this problem is dramatically reduced when they return to nighttime sleep. The sleep log data indicated a relatively high level of satisfaction with sleep quality, independent of shift worked or whether the sleep occurred during a workday. Generally, those dispatchers working the evening shift reported being more satisfied with their sleep quality than those working the day shift who, in turn, were more satisfied than those working the night shift. This is consistent with the shiftwork literature. The Naval Psychiatric Research Unit (NPRU) mood scores from the sleep logs do not appear to indicate any chronic sleep deprivation from these dispatchers. Measures of sleep efficiency and sleep fragmentation, derived from the actigraphy data, confirmed the quality of the dispatcher's sleep. Not surprisingly, the night workers were the most likely to use split sleep and naps to obtain adequate rest.

The experiences of this study suggest some modifications and enhancements to both the measures and protocol for future studies of dispatcher workload, stress and fatigue. First, future workload measures and protocols should try to capture the cognitive aspect of the dispatcher's work. With respect to a physiological measure of stress, salivary melatonin and ambulatory stress monitoring techniques should be considered. Regardless of the measures used to collect data during a dispatcher's work, more frequent data collection is necessary to capture the variation and short-term fluctuations in workload and stress that are inherent in the dispatcher's job. However, data collection more frequent than every 2 hr is probably not feasible in the workplace and a dispatching simulator would most likely be required. Further, it is recommended that workload, stress and fatigue each be explored separately first, before interactions among them are examined. Separate studies will enable a more in-depth understanding of each factor.

This field study was the first attempt to explore railroad dispatcher workload, stress and fatigue. Given that these data were collected from a small, non-randomly selected sample of dispatchers from two dispatching centers, the results should be carefully interpreted. This information should be used to gain insight into some aspects of the dispatcher's job and work environment and to identify areas that could benefit from further research. The findings and experiences of this study suggest the following issues for further research:

- Development of measures of dispatcher performance to correlate with workload, stress and fatigue.

- Development of a comprehensive demographic and health profile of the dispatcher workforce in the United States.

- Assessment of the value of planned rest breaks for reducing stress and fatigue in the dispatching environment.

- Development of a loss of alertness model for dispatchers.

- Measures and models of dispatcher workload that account for both the duration and intensity of workload over time.

- Assessment of the effect that Positive Train Control will have on dispatcher situation awareness, workload, stress and fatigue.

1. INTRODUCTION

1.1 Background

In 1987-88, the Federal Railroad Administration (FRA) undertook the National Train Dispatcher Safety Assessment. A number of factors led the FRA to initiate this study. The railroad industry's adoption of new dispatching technology, changes in operating rules and methods of operation, and railroad industry restructuring all had potential safety consequences. In addition, the FRA was concerned that excessive workloads and increased occupational stress could result from any of these factors.

During the course of the safety assessment, the FRA audited 125 dispatching offices and observed over 1,000 railroad dispatchers across the country. The FRA found evidence of occupational stress but felt that they did not have the expertise necessary to properly measure and evaluate the situation. According to the FRA findings, sources of stress appeared to be "frequent or occasional work overloads, ambiguous operating rules and instructions, the substantial safety responsibilities inherent in these positions, and on-time and maintenance requirements." In addition, the FRA observed a real or presumed lack of job security at some locations (FRA, 1990, p.9).

The FRA also collected data on the number of trains handled and authorities issued by individual dispatchers over the course of a shift. The FRA determined this to be an imprecise method of measuring dispatcher workload since it did not take into account the varied tasks that a dispatcher must perform to move a train across the assigned territory. In addition, this method captured neither the periods when the pace of work accelerates requiring the dispatcher to juggle several tasks and make rapid decisions nor the variability of workload within short time periods. The FRA concluded, "In order for useful data to be gathered, a system needs to be developed which could document the dispatchers' mental estimates of what is required to perform all individual tasks involved in the dispatching district. Parameters for measurement need to be established to assure workloads are indeed being measured. These parameters must determine how methods of operation, communication requirements and capabilities, control machines, computers and extraneous duties affect workloads" (FRA, 1990, p.11).

The subsequent Train Dispatchers Follow-up Review in 1993 underscored the observations and recommendations of the first dispatcher assessment with regard to stress and workload. The FRA again found evidence of occupational stress. Sources of stress appeared to be work overloads due to fluctuating traffic levels, coordinating maintenance of way work order authorities with high train movement periods, ambiguous operating rules, the substantial safety responsibilities inherent in the dispatcher's job, and the need to balance on-time train performance with maintenance requirements. While the FRA was able to identify important aspects of workload, they were not able to synthesize all of the disparate elements of workload into a coherent picture that adequately characterized the variability and complexity of the dispatcher's work. The FRA concluded that adequate methods for evaluation of both stress and dispatcher workload should be developed under a separate effort using experts in the health and human factors fields (FRA, 1995).

The FRA's interest in dispatcher workload, stress and fatigue is based on consideration of overall railroad operational safety. However, the FRA also recognizes the potential negative impact of these factors on employee health and well-being. Among other health problems, stress has been linked to heart disease, hypertension, various psychological disorders, and substance abuse (Weiten, 1992). In addition, the fatigue produced by work overload, night work, and shift rotation can have ill effects on sleep patterns, mood, and mental processes (Moore-Ede and Richardson, 1985). Maximizing employee physical and psychological health and well-being requires identification and evaluation of the features of the job, workplace, and organization that may erode the dispatcher's resources for meeting the demands of those tasks that place a great burden on attention, memory, and decision making capabilities.

1.2 Purpose

The purpose of the research described in this report was to understand today's dispatching environment and its associated levels of workload, occupational stress and fatigue. The project had the following goals:

- 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.

In addition to providing information to the FRA, this report is designed to serve as a resource to railroad industry officials interested in examining levels of workload, stress and fatigue among their dispatchers.

1.3 Scope

Field data collection was limited to two dispatching centers, one servicing a predominantly passenger operation and the other a predominantly freight service. Since these two sites represent a limited sample of railroad dispatching operations in the United States, the results of this study do not necessarily characterize all dispatching environments. The study addressed aggregate levels of dispatcher workload, stress and fatigue and the interaction among these factors. The study did not address the dispatcher's performance or efficiency, and non-job sources of stress were only considered to a limited extent. For example, differences in the three factors by years of dispatching experience, type of operation, age and work schedule were examined. As this was a study of aggregate levels, results for individual participants are not reported.

1.4 Overall Approach

The research was conducted *in situ*, that is, in the dispatcher's workplace. Assuring that the data collection did not interfere with the dispatcher's ability to perform his/her job was a major consideration in designing the study.

The overall study approach involved three phases:

- Phase 1: Selection of workload, stress and fatigue measures.
- Phase 2: Pilot study.
- Phase 3: Field study.

In Phase 1, candidate methods for measuring railroad dispatcher workload, stress, and fatigue were identified. A pilot study was then designed to test and evaluate the selected measures in the dispatcher's workplace and to ensure that the selected measures were appropriate to that environment, easy to use or administer, and did not interfere with the dispatcher's work. Based on the results from the pilot study, a full two-week field study was conducted at two railroads, one freight and one passenger operation.

1.5 Organization of the Report

This report presents the methodology and results of all three phases of this research project. Section 2 describes the nature of the dispatcher's job, the technology used in dispatching and the potential sources and consequences of dispatcher workload, stress and fatigue. Section 3 describes the data collection instruments and the rationale for their selection and Section 4 presents the study protocol. The study results are presented in Section 5. Finally, Section 6 assesses the methodologies used in the study and presents the study's key findings and recommendations. Appendix A contains information helpful in understanding the Actiwatch data. Appendix B contains copies of the forms used in the field study. Appendix C summarizes the results of the pilot study conducted to evaluate the procedures and instruments for field data collection, and Appendices D and E contain detailed data that support the results of Section 5.

2. THE NATURE OF RAILROAD DISPATCHING

Familiarity with the nature of the job of the railroad dispatcher, the dispatcher's workplace and the technology used by the dispatcher is a prerequisite to understanding the methods and results of this study. This section provides this prerequisite information along with some background from other work on sources of workload, stress and fatigue.

2.1 The Railroad Dispatcher's Tasks

The railroad dispatcher is responsible for the safe, efficient and economical movement of trains and other railway vehicles over the railroad, as well as for the protection of those who work on the railroad. The job of the railroad dispatcher consists of four basic job functions:

- Planning.
- Controlling track use.
- Managing unplanned and emergency events.
- Record keeping and report writing.

The job requires the dispatcher to issue, monitor, and cancel track usage authorizations according to specific railroad operating rules and procedures. The dispatcher also operates signals, switches and bridges, communicates with train and maintenance of way (MOW) crews and other dispatchers, schedules MOW work, responds to unplanned and emergency events, and performs administrative and clerical duties. As much as 75 percent of the dispatcher's shift may be spent communicating on the radio or telephone, leaving little time for accomplishing other duties (Vanderhorst, 1990). In this analysis of dispatching on the Burlington Northern Railroad, 38 percent of the communications topics addressed time and location, 14 percent addressed train information, 16 percent were the granting of movement authorities, 9 percent dealt with supplemental control, and 22 percent addressed special circumstances.

One of the first tasks the dispatcher performs at the beginning of the shift is to plan the known track moves on the territory for the duration of the shift. This strategy takes into account current traffic, expected traffic, the physical characteristics of the territory, train priorities (e.g., passenger versus freight), track and signal maintenance requirements, crew logistics (e.g., Hours of Service limitations), characteristics of train performance, and the presence of hazardous materials. Information from MOW crews, yard personnel, and dispatchers on adjoining territories also contribute to the dispatcher's actions.

Once this plan is in place, dispatchers spend the remainder of the shift trying to keep the trains moving while adjusting for the inevitable occurrence of delays and unforeseen events.

Even short delays may necessitate the reformulation of the entire plan. Indeed, the plan may have to be reformulated many times during the course of the shift as unplanned events transpire. The dispatcher is also responsible for monitoring train crew hours to ensure compliance with the Hours of Service Act. Finally, the dispatcher must also assume responsibility for railroad and non-railroad problems that are phoned in by railroad personnel, as well as the general public, and must be knowledgeable about the proper procedures for notifying the authorities in the event of an emergency, such as a hazardous materials spill. Although most dispatchers today rely heavily on computer-aided dispatching there is still significant record keeping and report writing that they must do.

2.2 Changes in the Dispatching Environment

The railroad dispatcher's responsibilities have remained the same for at least the last 25 years, but changes in the technology of dispatching and restructuring of the industry have resulted in a significantly different work environment. Each of the six basic job functions delineated in 1974 by Devoe is still relevant to today's dispatching environment and accurately describes the duties of a railroad dispatcher (Devoe, 1974). While the basic functions of the dispatcher have not changed, the past 20 years have seen significant technological and operational changes in the dispatching environment. The introduction of computer assisted railroad dispatching and communications has made it possible for dispatchers to control larger territories from much farther away. Changes in operating rules have allowed radio-transmitted directives to be used in place of traditional operator delivered train orders.

Concurrent with the increased reliance on computers, large centralized dispatching centers have evolved due to railroad mergers and consolidations. Today, dispatchers for the larger Class 1 railroads work in shifts around the clock in large centralized operations along with as many as 45 other dispatchers, and may control territories that are located over 1000 miles away. Changes in signal technology have led to a reduction in the use of tower operators and other field operations personnel, resulting in more direct dispatcher control over train movements, an increase in responsibilities, and an increase in the number of individual tasks involved in carrying out the same responsibilities. While some believe that these changes have made the dispatcher's job easier, others argue that they have led to increases in the dispatcher's workload and associated job stress and fatigue.

2.3 Dispatching and Train Control Technology

Current dispatching technology ranges from radio directives and paper forms, to almost "paperless" offices, where movement authorities and reports are completed using a computer-aided dispatching system (CADS). Typically, the larger the railroad, the more technology the dispatchers have available.

In computer-aided dispatching systems, train movement authorities, changes in switch and signal status, and other information about trains are entered into computers situated at each dispatcher's desk. Specific desk configurations vary by operation, but dispatchers likely have one or more computer screens and a keyboard at their desks, as well as a voice communications system. Typically, one or more computer screens present a schematic of the interlockings and control points of a territory for which the dispatcher is responsible, and over which the dispatcher has control, and shows track occupancy or other conditions of sections of track. Other screens may be used for data entry or information retrieval. In many instances, schematics of all of a railroad's territories are displayed in front of the dispatching office so that dispatchers are able to view their own territory as well as adjacent territories. Using computers at their desks, railroad dispatchers can change signals and switches, and enter and retrieve information about trains (e.g., train identification, locomotive power, train size and consist). Some computer-aided dispatching systems also record every keystroke and entry that is made by the railroad dispatcher for future review and analysis. Figure 1 depicts a state-of-the-art railroad dispatcher's console.

There is also a wide range of train control technologies that is currently being used, or being explored for use in the future. At one extreme, there are "dark" territories that do not contain any signalized systems. Trains are moved using hand-written or verbal movement authorities issued over voice radio by the railroad dispatcher. Before the advent of signalized systems and computer-aided dispatching, all dispatching was conducted in this manner, hence the term "paper railroad" to describe this manner of railroad dispatching. At the other end of the technology spectrum, there are currently several railroads that are working on a demonstration system of positive train control (PTC). PTC will ensure safety, and at the same time reduce the space between trains to increase throughput. Eventually, PTC will rely on digital data link communication, the Global Positioning System (GPS), onboard computers, and other advanced

Figure 1. Railroad dispatcher's console with CADS

technologies, and will likely change the way in which trains are dispatched. Among other changes, PTC will increase railroad dispatchers' reliance on computers to dispatch trains. Perhaps the most significant change is that under PTC the dispatcher's job will shift from actively controlling track use to monitoring its use.

2.4 Sources and Consequences of Dispatcher Workload

Workload may be defined as the interaction between the demands of a given task and the ability of the operator to meet those demands. It is a multidimensional concept, which may include elements of time pressure, pace of work, task difficulty and complexity, perceived control over work, and the level of effort and frustration associated with task performance. The dispatcher's workload is a mental rather than physical workload.

2.4.1 Workload and Performance

Workload can affect performance in a number of ways. First, high workload can degrade performance when the demands of the task exceed the resources that the operator is able to devote to it. Examples where this might happen include jobs that involve multiple simultaneous tasks, and tasks that involve rapid decision making and problem solving. Low workload has also been shown to negatively impact performance on various tasks. In this case a task may be so simple or repetitive that the operator's attention is not optimally engaged, and subsequently performance suffers. Examples may include sorting and low target frequency vigilance tasks and routine paperwork. Finally, performance suffers the most as a result of surges in workload, whether or not they are predictable. A predictable surge may result from rush hour traffic while an unpredictable surge could result from an accident or other emergency situation.

2.4.2 Contributors to Dispatcher Workload

Due to the complexity of the dispatcher's job, as described above, there are numerous contributors to workload. In its second dispatcher audit (FRA, 1995), the FRA identified the following factors as contributors to workload:

- Number of trains handled.
- Number of authorities issued.
- Number of control points/interlockings in a territory.
- Number, type and effectiveness of communications devices.
- Methods of operation.
- Total track miles in the territory.
- Administrative duties and paperwork.

These factors, along with others such as coping with an emergency situation, provided the basis for the exploration of dispatcher workload in the present study.

2.5 Sources and Consequences of Dispatcher Stress

Like workload, stress is a complex construct with many causes and manifestations. Stress may result from the individual's inability to cope with events or a situation. The subjective or cognitive experience of stress may include feelings of anxiety, anger, fear, helplessness, frustration, irritability, and depression. These feelings arise from the way an individual perceives and interprets events in his or her environment. As this is a very idiosyncratic process, events experienced as stressful for one person may be routine for another.

Stress occurs for many people in their work lives and can be categorized according to the source of stress. First, the actual demands of a task; that is pace, work volume and deadlines, may create stress within an individual. This illustrates the relationship between stress and workload. Second, environmental features of the workplace can cause stress. These features may include such items as lighting, air quality, and temperature. Finally, organizational issues such as work climate, work scheduling, conflict, change, and pressure may cause the individual a considerable amount of stress.

The follow-up FRA audit (FRA, 1995) of railroad dispatchers identified numerous sources of dispatcher stress. These were the following:

- Work overload and surges in workload.
- Juggling maintenance and traffic requirements.
- Ambiguous operating rules and procedures.
- Inconsistent application of rules between areas.
- Safety responsibilities.
- Threat of relocation.
- Radio frequency interference.
- Concerns about training.
- Changing technology.

As was the case with the workload issues, the FRA expressed reservations concerning the audit's ability to adequately inventory and describe all of the sources of dispatcher job stress.

2.5.1 Stress and Performance

Stress is associated with performance degradation and error production on the job. It causes a narrowing of attention, such that an operator may become fixated on one aspect of a problem and ignore other sources of information. Stress may also fragment attention, causing operators to search for information and solutions in a disorganized, unsystematic way. Becoming distracted as a result of stress may lead to procedural errors such as performing the wrong response or failing to respond to an important event; that is, a breakdown of the dispatcher's situational awareness. Other effects on cognitive functioning include poor decision making and problem solving ability, disorganization, forgetfulness, and distractibility.

2.5.2 Stress and Health

In addition to degrading performance, exposure to unremitting stress in the workplace can have a negative impact on employee health and well-being. Indeed, chronic work stress can lead to burnout, a condition in which the individual becomes so physically and psychologically exhausted that he or she is no longer able to function effectively on the job (Dell'Erba, Venturi, Rizzo, Porcu and Pancheri, 1994). Although it is difficult to definitively prove direct causal relationships, occupational stress has been linked to the following physiological and psychological effects (Weiten, 1992; Moore-Ede and Richardson, 1985; Knauth and Rutenfranz, 1987):

Psychomatic Disorders

- Heart disease.
- Hypertension.
- Skin problems.
- Gastrointestinal problems.
- Sleep disorders/circadian desynchronization.
- Asthma.
- Immune system suppression (infections, cancer, autoimmune diseases).
- Musculoskeletal pain/discomfort.
- Headaches.
- Impotence.

Mood-State Changes

- Anxiety.
- Depression (sadness, helplessness, loss of hope).
- Excessive worry.
- Anger/hostility.
- Irritability.
- Loss of motivation.
- Burnout (physical, mental, emotional exhaustion).

Psychosocial Effects

- Alienation from family, friends.
- Alcohol and drug abuse.
- Workplace violence.
- Domestic violence.

2.5.3 Stress of Dispatching

There is little research pertaining to the effects of stress on railroad dispatcher health. The research that does exist, however, suggests that occupational stress is a likely risk factor for ill health and decreased longevity among the dispatcher population. Most of the existing data have been compiled by the American Train Dispatchers Department (ATDD) of the Brotherhood of Locomotive Engineers and published in the Devoe Report (Devoe, 1974). Those studies are briefly described here, along with a more recent investigation.

A 10-month study conducted in 1929-30 showed an abnormally high rate of heart, blood, kidney, and anxiety problems among 165 train dispatchers. The average age of dispatcher death was, at the time, 50 years, with a very high proportion due to cardiovascular disease. Research dating from the 1940s and 1950s shows similar trends. McCord (McCord, 1948) found that the average lifespan of a dispatcher was 50.1 years, as compared to 65.9 years for age matched white males. Of the dispatchers studied, 81percent had diseased hearts and blood vessels. A study from the 1950s showed that 50 percent of dispatcher deaths were due to heart disease and another 20 percent involved diseases of the blood vessels (Devoe, 1974).

More recently, Menotti and Seccareccia (Menotti and Seccareccia, 1985) conducted a five-year mortality study of nearly 100,000 Italian railroad workers, aged 40 to 59 years. Workers were classified according to the level of physical activity (low, moderate, high) and responsibility (low, moderate, high) involved in their positions. They defined "responsibility" as the degree to which employees were accountable for loss of life, injury, and economic losses. The researchers found that railroad workers with low activity/high responsibility jobs were at greater risk than other groups for myocardial infarction. Although subjects were not classified by job title in this study, the dispatcher's job clearly fits the description of one that is both sedentary and high in responsibility. A five-year study of railroad personnel conducted in the 1970s is also worth noting with respect to stress and heart disease. Researchers noted an exceptional number of smokers among train dispatchers, as well as a very high rate of death due to coronary heart disease (Devoe, 1974). This finding indicates that smoking behavior may serve as a moderator variable in the relationship between train dispatcher stress and subsequent illness. Finally, cardiac problems have been associated with working shifts, and particularly the night and rotating shifts. Other problems associated with shiftwork include sleep disorders and attendant fatigue, gastrointestinal problems, alcohol and drug abuse, social isolation, and disorders of mood (Moore-Ede and Richardson, 1985).

2.6 Sources and Consequences of Dispatcher Fatigue

Mental and physical fatigue may also interfere with the train dispatcher's work. Although the job does not require hard physical labor, it is important to keep in mind that shiftwork and sleep deprivation may cause physical fatigue, and that this in turn is likely to affect the dispatcher's level of mental fatigue and alertness. Mental fatigue can also accumulate as a result of time on task, large work volumes, rapid information processing and decision-making, and responding to problems such as emergency events.

2.6.1 Mental Fatigue and Performance

The symptoms of mental fatigue involve loss of alertness, feelings of sleepiness, lack of energy, weariness, and exhaustion. The effects of fatigue on performance include slower reactions, poor concentration and forgetfulness, complacency, and an increasing reluctance to expend any effort in task performance. Fatigue is also associated with on-the-job sleepiness and microsleeps (Weiten, 1992). When an individual is microsleeping, s/he appears to be awake (sitting upright, eyes open), but an electroencephalogram (EEG) would indicate that the person is actually in a light sleep.

2.6.2 Fatigue and Shiftwork

There are two major contributors to fatigue among train dispatchers. First, staffing shortages result in overwork. A shortage of relief employees results in the dispatcher having to work on normal rest days. In addition, where staff shortages exist, it is not uncommon for the dispatcher to work for periods exceeding what is allowed under the Hours of Service Act (FRA, 1995). Under the Hours of Service Act, dispatchers may not remain on duty for more than 9 hr in any 24 hr period where two or more shifts are employed. Where one shift is employed, the duty period must not exceed 12 hr during any 24 hr period.

A second source of fatigue involves the shiftwork system. This is of particular concern for nightworkers and those who work rotating shifts. Rotating schedules can have adverse effects because the body's circadian rhythms do not have time to adjust to any single schedule. Even when one works third shift consistently, there is a long period of adaptation required for re-entraining physiological functions (Knauth and Rutenfranz, 1987). In fact there is some question as to whether complete adaptation to night work is even possible. The degree of adaptation depends upon both the individual and the length of time spent on the shift. One longitudinal study, however, shows incomplete adaptation even after three years working the night shift (Dahlgren, 1981). Thus, at the very least, it is likely that night shift employees will experience long periods of fatigue and loss of alertness while adapting to the night shift schedule. In addition to the fact that daytime noise will often interfere with a night shift employee's sleep, the nightworker may also lose sleep because of family and personal obligations that must be attended to during the day. The conflict between the need to sleep and the desire to spend time with family and friends can also become a significant source of stress in an individual's life (Moore-Ede and Richardson, 1985).

3. DATA COLLECTION INSTRUMENTS

The overall approach to identifying and selecting candidate measures of workload, stress, and fatigue was threefold. First, related studies in other fields, such as air traffic control (ATC) and vessel traffic services, were reviewed. Next, site visits to local dispatching centers provided an assessment of what data collection methods would be appropriate for this work environment. Finally, a focus group interview with train dispatchers provided insights on their concerns regarding workload, job-related stress, and sources of fatigue as well as the suitability of candidate data collection instruments. Based on the information gathered from the literature review, site visits and focus groups, candidate data collection instruments and methodologies were identified and selected.

3.1 Data Sources from Related Research

Initially, the study team investigated a number of candidate measures to be used in this study, selecting those instruments it felt most appropriate for measuring dispatcher workload, stress, and fatigue. Sources of information included recent literature, test banks, personal communications with other researchers in the fields of interest, and:

- Reviews of recent stress and shiftwork literature.
- Workload and stress measures used with ATCs.
- Fatigue measures used for truck drivers and locomotive engineers.
- Stress and workload surveys used by NIOSH.
- Stress questionnaires contained in several test banks.

3.2 Site Visits

Site visits to several dispatching centers provided valuable insights about the job of dispatchers as well as the environment in which they work. During these site visits potentially stressful and workload-inducing elements of the dispatcher’s job, the ergonomics of their workstations, and organizational issues that may serve as additional occupational stressors were observed.

3.3 Focus Group

A focus group interview was conducted with railroad dispatchers from two railroads, as well as representatives from the ATDD, the dispatchers’ labor union. The goals of the focus group were to:

- Provide feedback on several of the survey instruments that were currently under consideration.

- Comment on the acceptability of the candidate physiological measures (i.e., saliva collection, actigraphy).

- Identify additional sources of workplace stress, workload, and fatigue.

- Identify potential concerns of study participants.

Participants confirmed that the volume and pace of the dispatcher's work would not allow for the use of intrusive and/or time-consuming measures. It was clear that study participants would not have time to take breaks during the workday in order to fill out multiple and/or lengthy test instruments. One concern was that study participants might not be able to complete the test instruments at the same time each day, due to variations in daily traffic volume, and the possibility of schedule changes and unplanned events. More importantly, the focus group participants expressed concern that the test instruments would distract them from their work and that they could make an error as a result. This was a particular concern since dispatchers can be held personally liable for any accident or emergency situation that occurs on their watch and that is due to dispatcher error. Consequently, all of the measures used in the study had to be simple, brief, and unobtrusive. Due to liability considerations and the fact that the study was conducted in the workplace, videotaping of participants was not feasible.

3.4 Selection Criteria for the Data Collection Methodologies

The feedback and information provided by focus group participants necessitated revisions to the original study plan. The formal tradeoff study that was originally proposed to compare the different data collection methods was precluded, since it became clear that the nature of the dispatcher's work and workplace would necessarily dictate the measures ultimately used. Specifically, dispatcher acceptance became the single criterion for selection, although such acceptance is dependent upon several factors: ease of use, time to administer, and lack of interference with duties.

Given these factors, any instruments or methods that would interfere with the employee's concentration, communications, and work processes were not feasible for this study. Not only would multiple and/or lengthy test instruments administered during the workday be unacceptable to potential volunteers, but perhaps to railroad management as well. In this regard, it seemed likely that the use of a large battery of test instruments would alienate candidate railroads and make it difficult to secure a test site.

3.5 Survey Instruments

Two survey instruments were developed. The background survey instrument was designed to capture demographic and work-related information for each participant. The debriefing survey instrument was designed to collect information from the participants about the conduct of the study as well as to solicit opinions regarding desired workplace improvements.

3.5.1 Background Survey Instrument

A background survey instrument was developed for administration to the study participants prior to the onset of the actual field data collection. Focus group participants reviewed a first draft of this survey instrument. The survey was modified in accordance with their concerns and feedback. An interim meeting between the FRA, industry representatives and union officials was also used to evaluate and augment the survey. The resulting background survey instrument was comprised of 10 subsections and took approximately one-half hour to complete. The subsection topics focus on sources of dispatcher workload, job stress, and fatigue, as well as information regarding employee health, work scheduling, sleep habits, quality of life issues, demographic data, job satisfaction, and work climate.

The purpose of the background survey instrument was twofold. First, it provided normative data on this worker population. These data, in turn, could then be used to characterize the population by both allowing the creation of an overall stress profile of the railroad dispatcher, and having a base of information from which comparisons can be made to the general U.S. population. As noted in subsection 2.5, stress that is associated with task performance on the job is not the only source of stress in an individual's life, and stress from different sources will tend to interact in a complex manner. For example, other job-related stressors might involve various organizational problems—poor relationships with colleagues or supervisors, issues related to shift work and work scheduling, and stress that occurs in response to organizational or technological changes. Stress that is related to personal, family, social, and financial matters is also likely to contribute to on-the-job stress. Thus, data from the background survey instrument was designed to address: 1) what factors, in addition to job demands, are creating stress in the lives of dispatchers, and 2) what, if any, stress-related health symptoms are currently being manifested.

3.5.2 Debriefing Survey Instrument

The debriefing survey instrument served a dual role. First it solicited participants' opinions and suggestions concerning the methods that were used to collect information for the study. In terms of the pilot study, this information provided important input to the development of the final protocol used in the full field study while survey information from the full field study will be helpful in future studies of this nature. Second, this survey instrument was used to collect final thoughts and information on dispatcher issues not otherwise addressed in the study protocol. For example, this survey instrument asked for opinions regarding desired improvements to the job/workplace, top five contributors to work stress, how they defined "workload" and when their last days off were prior to the start of the study.

3.6 Measures of Workload

This study used three distinct approaches to measure workload: objective workload, subjective workload and Task Analysis Workload. The following sections describe each of these approaches in detail.

3.6.1 Objective Workload

Objective workload refers to the task demands of a particular job, independent of the individual's abilities to meet those demands, independent of the human operator. Two approaches to objective workload were considered, an observational technique employing a software tool and paper records.

3.6.1.1 NIMS Data Collector

The NIMS (NAS Infrastructure Management System) data collector is a software package consisting of a simple user interface and integrated database. It was designed to gather task frequency and duration information on NAS (National Airspace System) Airways Facilities personnel as they perform their job duties (Nadler, Haines, Bonin and Grossberg, 1997). Modifications to the current interface and data collection parameters were made to reflect the unique needs of observing the railroad dispatching environment and to meet the needs of the current project.

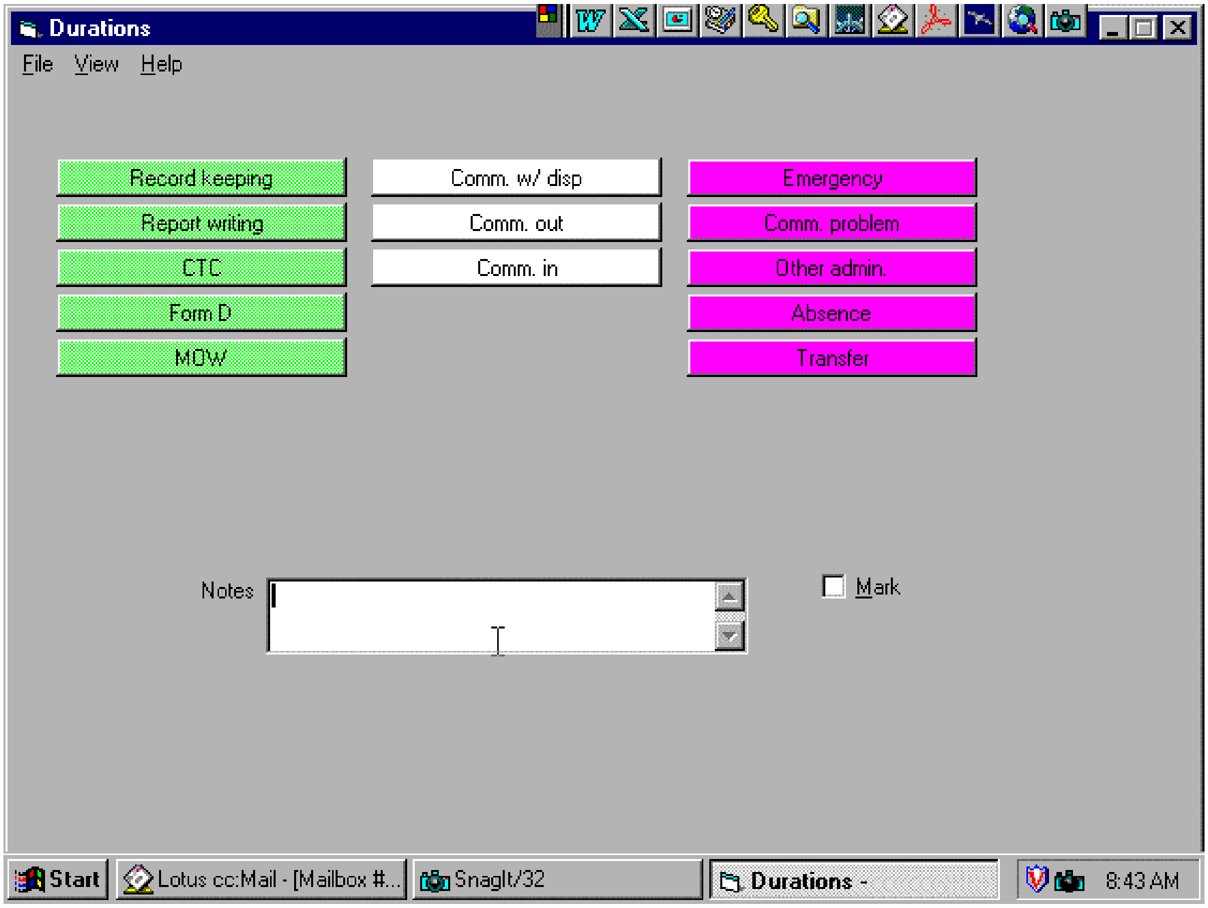
The NIMS data collector required a trained research associate to observe the railroad dispatcher at work and, using a laptop computer, record data corresponding to specific dispatcher tasks. (Figure 2 displays the NIMS data entry screen.) Observations focused on activities that take place within several broad categories of dispatcher responsibility, including communications, paperwork, monitor watching, and data entry.

The dispatcher's job is broken down into distinct tasks that are represented on the computer's interface. An observer must watch a particular dispatcher and make entries into the computer system that correspond to the start and stop times of specific tasks. This is done in real time. The program yields a running record of all of the dispatcher's activities during the observation period and includes information on the duration, sequencing, and frequency of each task.

Following a visit to a railroad dispatching center, a preliminary list of dispatcher tasks was developed. This list was refined based on input from a subject matter expert (SME) and a review of Devoe (Devoe, 1974). The resulting 13 dispatcher activities (see Table 2) were validated during a visit to a local dispatching center to ensure that no activities were missing and that the activities included were representative of the dispatcher's workload.

3.6.1.2 Dispatcher Records

Due to the inherent difficulties of using the NIMS data collector in the dispatching environment, another measure of objective workload had to be identified. Paper-based records of the various dispatcher tasks are not time-sensitive since they can be recorded after the fact. In addition, they provide information similar to what was collected with the NIMS data collection system. Data collection during the pilot study indicated that this type of information was relatively easy to collect and appeared to be a valid reflection of workload. Specific workload variables and the source of the data are listed in Table 3. These data were collected in 2 hr intervals after start of shift.



510-FRA-00045-3

Figure 2. NIMS data entry screen (NIMS data entry screen.doc)

Table 2. Dispatcher activities

1.	Record keeping and record review.
2.	Report writing.
3.	CTC entry.
4.	Issue/cancel Form Ds and verbal train movement authorities.
5.	Issue MOW protection/roadway worker protection.
6.	Communicate with other dispatchers and Chief dispatcher.
7.	Dispatcher-initiated communications.
8.	Externally-initiated communications.
9.	Respond to unplanned and emergency events.
10.	Experience communication problems.
11.	Other administrative duties.
12.	Non-operational absence.
13.	Transferring responsibility to or from a desk.

Table 3. Sources of objective workload data

Measure	Data Source
Number of trains handled	Subjective rating forms
Number of other track users handled	Subjective rating forms
Number of Form D's issued/cancelled	Archived completed forms
Number of Route Blocks issued/cancelled	Archived computer printouts
CETC keypress information*	Archived computer printouts

*Data not collected due to dispatching computer technical problems.

3.6.2 Task Analysis Workload Measure

A number of factors make the task of assessing the workload of dispatchers difficult. First, many of the tasks involve cognitive activity that is not directly observable. Second, the timeframe of task performance is often unpredictable due to the delays that may exist between event onset and the initiation of a response. Finally, assessment is aggravated by the fact that several tasks are often performed concurrently during a given time period. Given these problems, an observational methodology and analytical technique based on the Task Analysis Workload (TAWL) was developed and explored for this study. TAWL, developed by the Army Research Institute, was originally intended for the assessment of the workload of military helicopter flight crews (Bierbaum, 1990). The TAWL was modified to suit the dispatching operational environment and is referred to as the mTAWL.

3.6.2.1 Description

The mTAWL is a task-oriented approach that assumes dispatchers perform multiple simultaneous tasks through time, and that these individual tasks may vary in their demand on the dispatcher's performance resources. The mTAWL treats workload as the sum of the difficulty of all concurrent tasks for each minute in an observation. Two dispatchers might 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 task difficulty 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 the auditory, visual, cognitive and psychomotor channels. The mTAWL method calculates workload by summing the loads for each of the individual channels across all tasks each minute of a dispatcher's shift.

The mTAWL approach assumes that a task, such as handling an unopposed train, places demands on a dispatcher that begin when the train comes under the dispatcher's jurisdiction and continue until the train leaves that jurisdiction. Given that there is no practical way to observe or calculate the exact, minute-by-minute impact of a task, the mTAWL assumes an average task resource load that is spread evenly over the duration of the task. This treatment of constant load distribution over the duration of a task is a reasonable alternative to attempting to infer what is truly going on in a dispatcher's mind on a moment to moment basis. Observed dispatcher actions that are directly related to an ongoing task are understood as consequences of the task rather than as something to be counted. While dispatcher actions that are related to ongoing tasks are not directly quantified in the mTAWL approach, they are used to make judgments about the nature of the task. For example, mTAWL makes a distinction between the handling of track occupants whose progress or position is in conflict with other occupants and track occupants that do not pose a conflict. Dispatcher radio and telephone transmissions, supported by train sheets or on-time station records, provide the basis for judging the opposed/unopposed status of a track occupant.

The mTAWL approach uses dispatchers familiar with a specific dispatching center to first define the tasks that make up a dispatcher's work and then to make judgments about the level of effort required from each of the four resource channels for each task. Trained personnel then observe dispatchers at the dispatching center over a period of time. After the data is collected, an independent rater applies the expert-developed criteria to the observer's records to create the mTAWL scores. Given the source of its criteria, and the dependence on rater judgment, the mTAWL analysis must be considered subjective despite its seemingly definitive graphic and numerical outputs. On the other hand, the original TAWL approach has been shown to correlate well with subjective self-rated measures of workload.

Adaptation of the TAWL approach for the present study of railroad dispatchers required defining a set of tasks that describes the dispatcher's job. Based on extensive observation of dispatchers at work and discussions with experienced dispatchers at each study site, a set of tasks was developed. The tasks fell into four conceptual areas:

- *Background tasks* are either continuous (e.g., monitoring track occupants) or scheduled at a time chosen by the dispatcher. These tasks do not require an immediate response and are not time critical.

- *Foreground tasks* are unanticipated, and are unrelated to current or previously anticipated train or work crew moves. Such tasks require an immediate response.

- *Anticipatory tasks* are preparatory actions and plans for managing trains or work crews whose approximate arrival time in the dispatcher's territory is known. The timeframe for this type of task begins when the dispatcher first becomes aware of the occupant, or in the case of a scheduled train, when the appropriate section of track is cleared to receive the train. The anticipatory period concludes when the expected track occupant enters the dispatcher's control or is no longer expected to arrive. Each anticipated track occupant constitutes a separate event.

- *Track management tasks* are those actions taken by the dispatcher to coordinate train movements and protect work crews on the dispatcher's territory. The timeframe of this type of task begins when the track occupant first enters the territory. It ends when the track occupant leaves the dispatcher's territory. Each track occupant constitutes a separate event.

Table 4 contains definitions of the specific tasks in each task area.

An experienced dispatcher identified the triggering events and concluding actions that mark the beginning and end points of each task in Table 4. Separately, based on expert dispatcher judgements, each task was scored on a scale from zero to seven according to its contribution to workload in each of the four information-processing channels: auditory, visual, cognitive and psychomotor. The mTAWL ultimately yields a graph that provides a means of identifying peak workload in one or more of the information-processing channels and the time(s) of the peak(s). (See subsection 3.6.2.3 for sample graphs.)

Table 4. Definition of mTAWL tasks

Task Area	Task Description
Background	• *Auditory monitoring* – listening for radio or telephone calls. • *Visual monitoring* – watching for unanticipated events on the informational displays. • *Background "phone"* – telephone or radio calls made during the course of work at times chosen by the dispatcher and not directly related to the anticipated or actual control of a specific train or work crew. • *Background "computer"* – computer or paper entries required during the shift that do not require an immediate action and can be performed at times chosen by the dispatcher.
Foreground	• *Foreground "phone"* - telephone or radio calls made or received during the course of work which required an immediate response and which were not directly related to previously anticipated actions or the current control of a specific train or work crew. • *Foreground "computer"* - computer or paper entries required during the shift that were not directly related to previously anticipated actions or the current control of a specific train or work crew but that had to be taken care of immediately.
Anticipatory	• *Anticipated unopposed* - the anticipated arrival of a track occupant whose progress toward the dispatcher's territory is expected to be unimpeded. • *Anticipated opposed* - the anticipated arrival of a track occupant whose progress to the dispatcher's territory was expected to be opposed by another track occupant or occupants.
Track Management	• *Track occupant unopposed* - a period of time during which a specific track occupant under the dispatcher's control is unimpeded. • *Track occupant opposed* - a period of time during which a specific track occupant under the dispatcher's control is opposed by another track occupant or occupants.

The application of the original TAWL to helicopter pilots assumed that, at any given period of time, when the workload score (summed across tasks) for a given information-processing channel exceeded the highest possible level of seven, then an overload condition existed. In flight scenarios, these periods are understood to be occasions when catastrophic errors are likely to occur. The dispatching environment was assumed to be a similar environment and initially a score of seven was used as an indicator of work overload.

3.6.2.2 Development of mTAWL Channel Loadings

The next step in applying mTAWL to a dispatching environment was to establish the channel loadings for each of the 10 tasks that was to be observed. Since every dispatching center is unique, a set of channel loadings must be determined for each dispatching location. The ratings can be established by assembling a small group of experienced dispatchers from the center where data will be collected. Assistant chief dispatchers and senior experienced dispatchers performed this task.

The constructs of dispatching tasks and task channel loads were first explained to the expert group. Then the list of tasks was presented and the expert dispatchers were invited to suggest additional tasks if they felt any were needed. In the present study, dispatchers from the freight operation agreed to add a "surcharge" to account for 1) additional work necessary for trains carrying excess dimension cars or 2) the failure of safety or signaling equipment. Effort in the anticipation or handling of track occupants under either of these conditions was thus increased. No additional tasks were identified for the passenger operation.

Next, the expert dispatchers were asked to independently choose levels for each channel. The final channel loads for each task were determined by taking the mode of the responses from the expert group for each channel. Where no modal value existed, the median value from the range of responses was used.

3.6.2.3 Analysis of mTAWL Data

The minute-by-minute channel loads may be represented in a number of ways. These are:

- Raw channel loads.
- Channel loads standardized for the individual.
- Channel loads standardized for all desks.
- Fifteen minute moving average channel loads.

Each approach is defined and discussed below.

The raw channel is computed by minute based on the tasks that the individual was performing at each minute and the weightings for those tasks. Figure 3 illustrates an example of the total channel load by minute. In this example, the auditory and visual channels overlap almost perfectly and thus the two traces appear as one. This first shift dispatcher's desk shows a low level of activity at the beginning of the shift and a period of high activity beginning

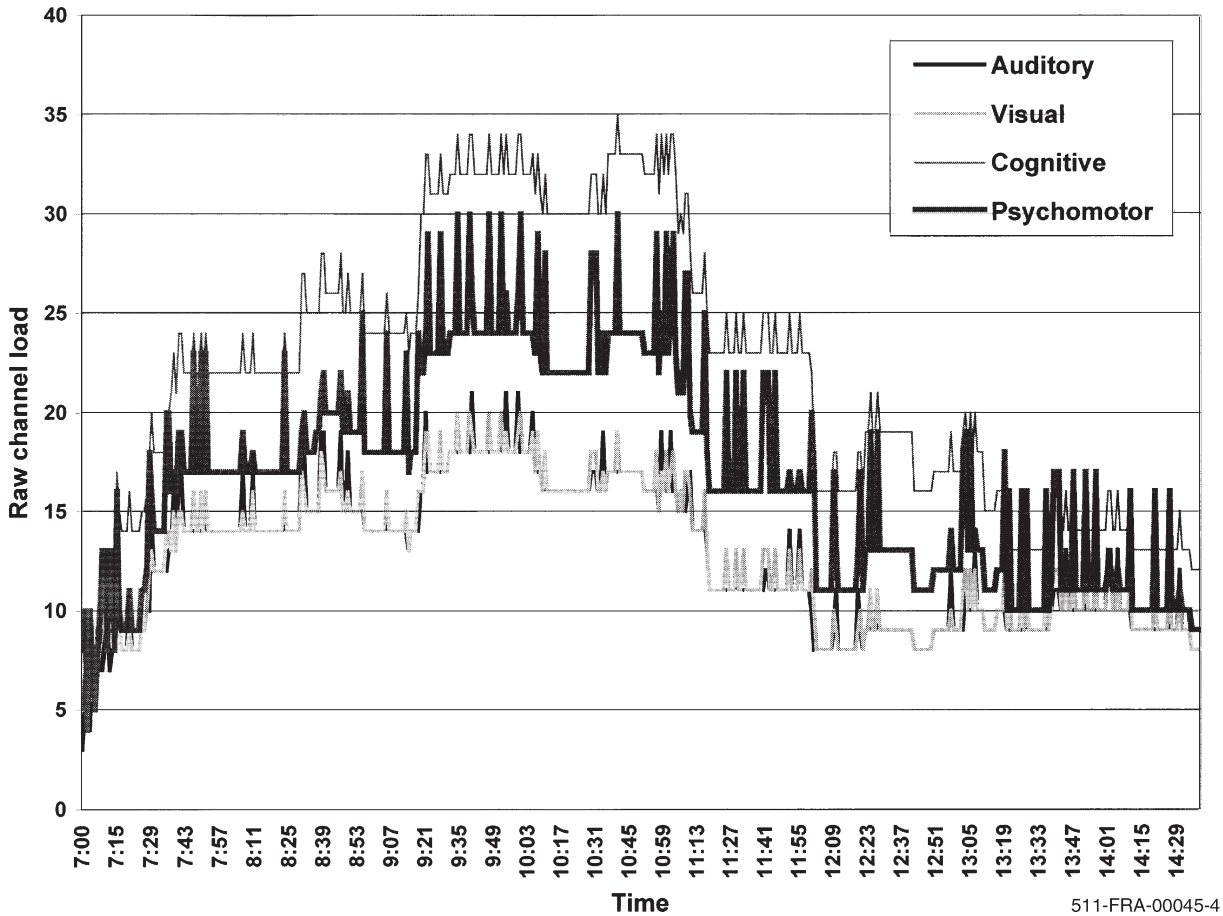


Figure 3. Raw channel loads

approximately 2 hr later. The observer's notes reveal that during the busy period this dispatcher handled a large number of track occupants. Due to the route blocking effects of track maintenance activities, most of the track usage authorities issued by the dispatcher during this time were for opposed moves. This graph also shows that the channel loads are highest for cognitive and psychomotor tasks. A review of the observers' notes shows a significant amount of time spent preparing paperwork and managing a large number of opposed track occupants that would account for the heavy cognitive and psychomotor loads.

While the mTAWL procedures employed in the pilot study produced results that were consistent with the concept of a channel load of seven indicating an overload condition, the channel loads were much higher in the full field study. There were a number of reasons why this may have occurred. The focus groups used to establish the task weightings assigned higher weights than were developed for the pilot study. Also, the mix and number of track users and the period of time that they remained on the dispatchers' territory also differed between the pilot test and the full field study. Since the concept of seven as a maximum load was not meaningful for the field study, an alternate means for identifying overload conditions was developed.

Channel loads were converted to standardized scores (Z-scores). The Z-score is computed from the following:

$$Z = \frac{\left(x_i - \overline{X}\right)}{\sqrt{\sum_{i=1}^{N}\left(x_i - \overline{X}\right)^2 / N}} \tag{1}$$

where

$\overline{X}$ = arithmetic average for a group of N scores, and
x_i = score for the i^{th} dispatcher

This transformation produces a distribution of scores with an average of zero and a characteristic that allows direct comparison between distributions. In addition, probabilistic statements can be made about individual scores based on their distance from zero. In many ways, this is more useful than arbitrarily selecting a number to represent overload. For example, the probability is that approximately 68 percent of all scores are expected to fall between Z-score values of +1.0 and –1.0 and approximately 95 percent of all scores will fall between Z-score values of +2.0 and –2.0. This means that a channel Z-score greater than 2.0 is an unusual event, likely to occur less than 5 percent of the time by chance alone.

The standardized channel Z-scores were computed in two ways. First, individual dispatcher Z-scores were calculated to show deviations in load relative to that dispatcher's individual work on a given shift. Figure 4 shows the data of Figure 3 transformed into Z-scores using the individual's own task channel loads from the shift. Zero represents the average load for this dispatcher on this shift. As can be seen in Figure 4, for the last 3 hr of the shift, the workload per channel is actually below average for this individual.

One distinction that is lost through standardizing the channel ratings is the absolute difference between the channels. The standardized Z-score treatment conceals the fact that the auditory and visual channels for this specific observation constituted a much lower demand than the cognitive and psychomotor channels.

While the self-comparative nature of the previous approach is informative, more is revealed if the workload figures can be compared across all desks at a dispatching center. Z-scores were computed as described above using Equation (1), but for this analysis deviations are relative to *all* 10 desks observed at the center. In Figure 5, the channel loads by minute for 10 different dispatcher shifts at the same center are used to standardize the channel loadings for the dispatchers' workload shown in the previous examples.

In Figure 5, zero now represents the average for each channel load for all 10 dispatchers at the dispatching center. In comparison to the previous example, it becomes apparent that the participant's channel loadings displayed in the graph were somewhat higher than those of the

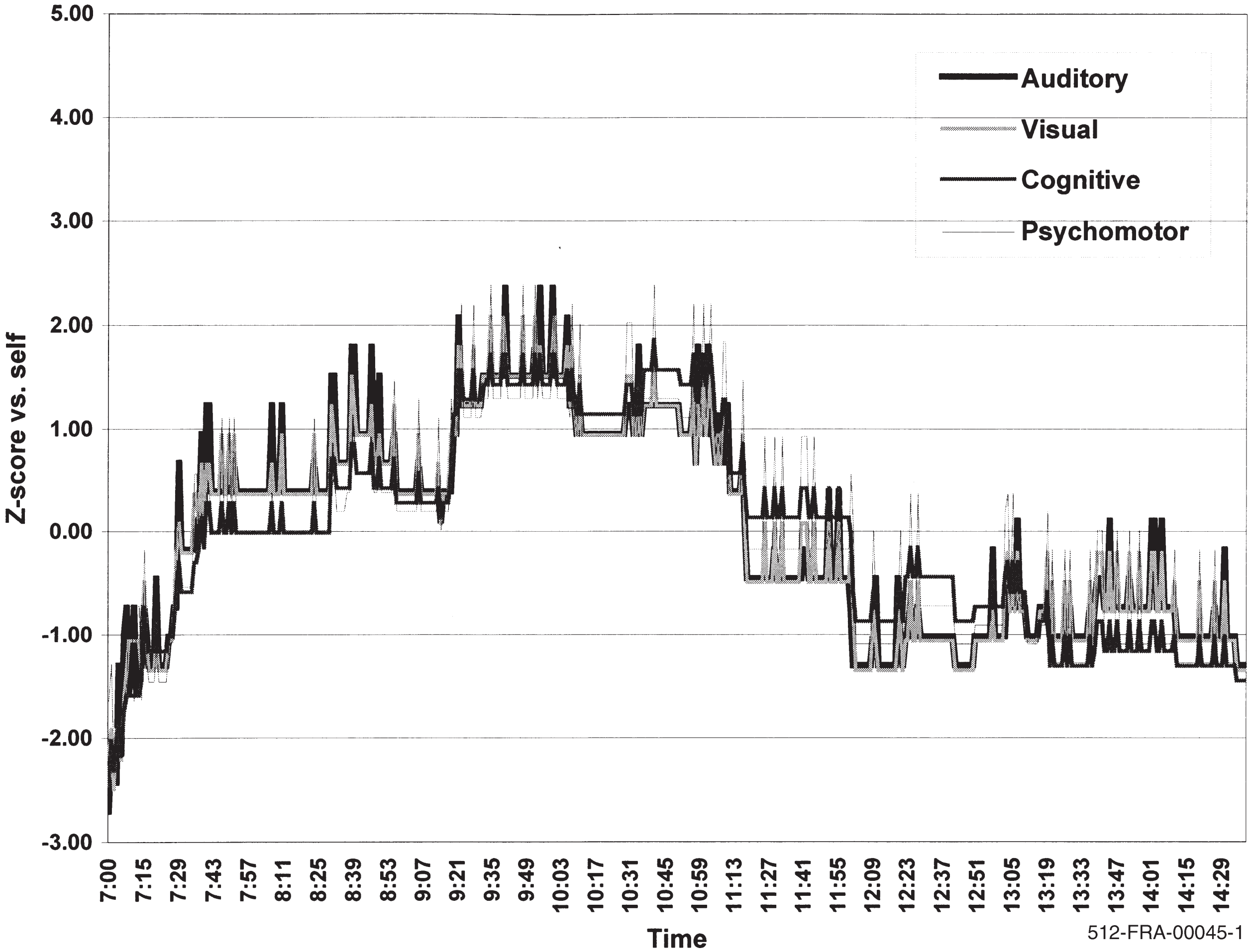


Figure 4. Channel data standardized against individual

group as a whole. The third and fourth hours of the shift were somewhat busy in comparison to the subject's own averages, but were extremely busy in comparison to the group as a whole. Similarly, the last 4 hr of the shift were a below average channel demand for the subject, but constitute an average load when compared to the group of 10 dispatchers.

Since the raw channel data as well as the standardized Z-scores many times do not allow for identification of a trend, a fourth method, moving average, was applied to the standardized scores. A 15 min moving average was applied and is shown in Figure 6.

The 15 min period was chosen because it was felt that a 15 min period of continuous high Z-scores indicated an overload condition.

The determination of what constitutes overload is, of course, subjective. One of the problems is that dispatchers vary in skill. A flow of work that might be an overload for one dispatcher may not be for another. Regardless, a standardized workload score, relative to the group of dispatchers, that exceeds ±2.0 for periods of 15 min or more constitutes a substantial deviation from average and certainly invites a careful review to determine the conditions that produced it.

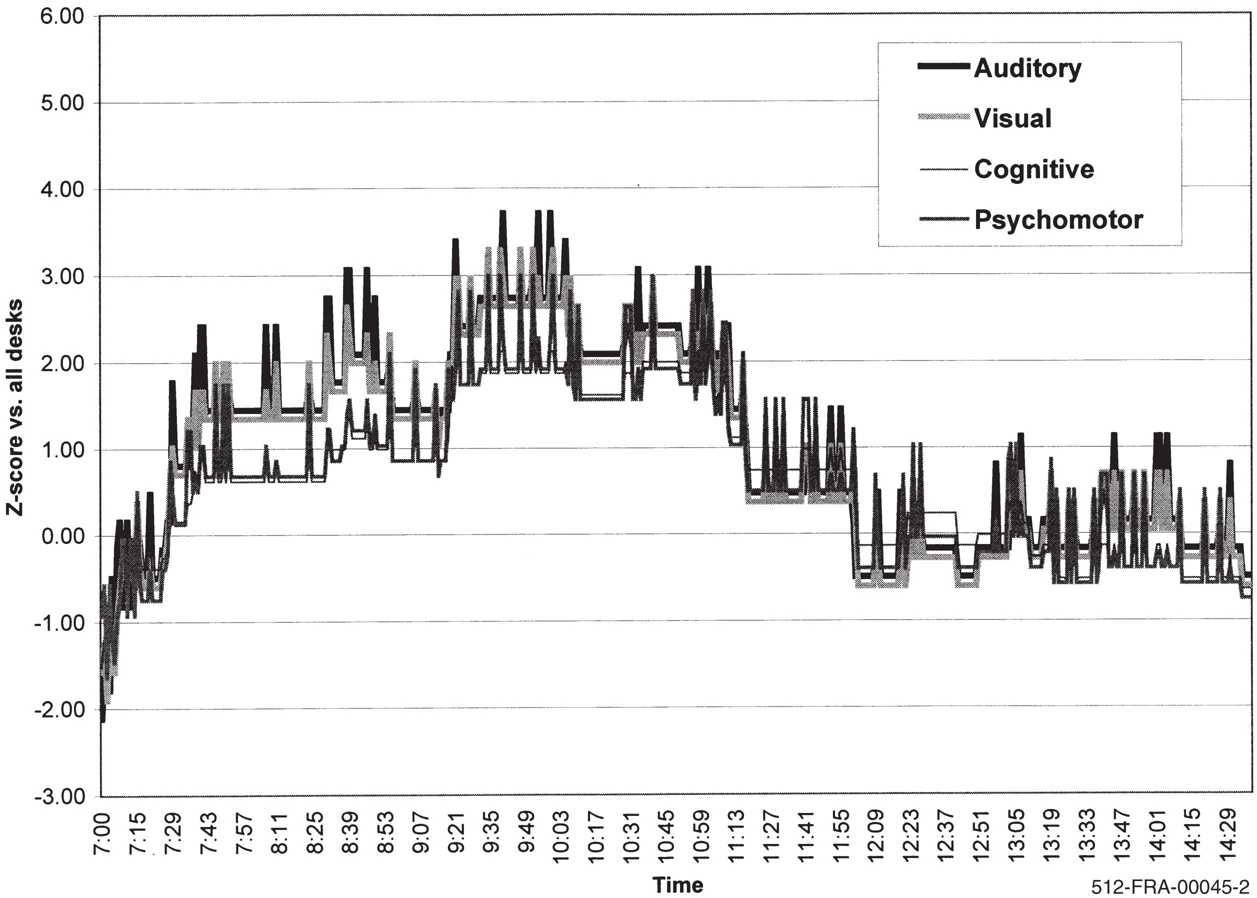


Figure 5. Channel scores standardized against center

3.6.3 Subjective Workload Measures

Subjective workload refers to the railroad dispatcher's perception of both: 1) the demands of their work in terms of difficulty, complexity and time pressure, and 2) the effort they need to expend to meet those demands. This construct is often assessed using self-report rating scales that are administered during or immediately following task performance. Such scales may be administered several times during a work period in order to document changes in perceptions of workload over time. There are several popular subjective workload techniques available, both unidimensional and multidimensional. Scales are defined as unidimensional if the operator is required to rate the *overall* workload of a task only. Multidimensional techniques are more diagnostic in the sense that operators are asked to rate different elements or dimensions of the workload, such as time pressure and mental demand.

The simplest unidimensional workload scale is the Overall Workload (OW) scale. It is a unidimensional 100 mm bipolar visual analog scale ranging from 0 to 100 (Hart and Staveland, 1988). Operators are typically asked to place a hash on the part of the scale corresponding to their perceived level of workload. Other unidimensional scales are more complex. For example,

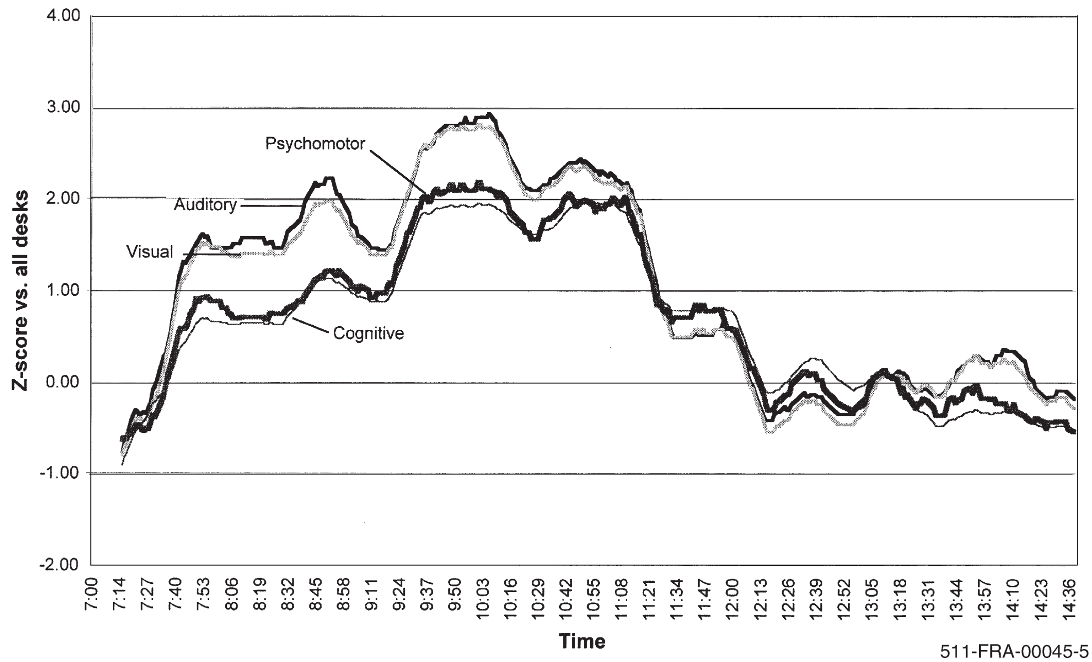


Figure 6. 15 min moving average channel load

the Modified Cooper-Harper Scale (MCH) requires operators to work through a multi-faceted decision tree in order to arrive at a global workload score (Cooper and Harper, 1969). The Bedford Workload Scale also utilizes a decision tree (Roscoe, 1987). The Bedford Scale, while designed for pilots, can be modified for use with other types of operator.

There are also a number of multidimensional workload scales. With the NASA Task Load Index (TLX), operators rate six sources of workload—mental, physical, and temporal demand, performance, effort, and frustration—each on a scale ranging from 0 to 100. Choosing the greater contributor to workload in pairwise comparisons produces a weight for each source of workload. Ultimately, this process yields a weighted rating for each of the six workload dimensions, as well as an overall, global workload rating. (Hart and Staveland, 1988). Another multidimensional workload rating technique that was investigated is the Subjective Workload Assessment Technique (SWAT; Reid and Nygren, 1988). SWAT asks operators to rate tasks on three dimensions: Time Load, Mental Effort, and Psychological Stress. The SWAT also entails a labor intensive sorting and weighting process to determine the relative contributions of each subscale to overall workload.

While the multidimensional workload instruments provide more specific, diagnostic information about the sources of an operator's workload, they are too time consuming and complex to administer to an on-the-job railroad dispatcher several times over the course of his or her shift. The decision tree methods can also be time intensive. It was determined that the best subjective workload technique to use in this particular operational setting is the OW in conjunction with the three NASA-TLX sub-scales that are most appropriate to dispatcher workload: Mental Demand, Temporal Demand, and Effort. These four scales simply required a single rating between 0 and 100. The NASA weighting process for the subscale items was not utilized. Figure 7 contains an example of the type of subjective scale used to rate these four factors; it employs the 100 mm visual analog technique.

3.7 Measures of Stress

This study used two distinct approaches to measure stress: salivary cortisol and subjective stress. The following sections describe each of these approaches in detail.

3.7.1 Salivary Cortisol

In addition to a subjective report of stress, options for taking a physiological stress measure periodically during the dispatcher's work shift were explored. Such measures include pulse rate, sustained contractions of the frontalis muscle (forehead), and cortisol secretion. The latter option appeared to be the most appropriate given the constraints of the data collection environment. Cortisol is a reliable stress marker that can be collected quickly and unobtrusively in subjects' saliva. For example, cortisol has been used in studies of air traffic controllers (Fibiger, Evans and Singer, 1986, p. 29-36).

Psychological and physical stress increase the activity of the hypothalamic-pituitary-adrenal (HPA) axis of the central nervous system, which results in release of corticotropin releasing hormone (CRH) and adreno-corticotrophic hormone (ACTH). These hormones subsequently give rise to an increase in the release of cortisol from the cortex of the adrenal gland, and consequently a rise in cortisol concentration in the blood, urine and saliva. Sampling of free cortisol in blood, the traditional method for controlled laboratory and clinical investigations, typically requires continuous sampling via indwelling venous catheters, fraction direction devices or techniques, and continuous monitoring of research subjects by skilled technicians.

By contrast, saliva collection can be done non-invasively and outside the constraints of a research laboratory. Study participants can quickly and easily learn to collect their own saliva

Figure 7. Example of subjective 100 mm visual analog scale

samples thus allowing the possibility of frequent on-the-job sampling. This type of sampling can be assessed at relatively low expense and without disrupting the regular routines of the participants. Most importantly, the sampling of salivary cortisol is a stress-free procedure. This is critical because both plasma and salivary cortisol are elevated transiently in response to stress, including the psychological stress generated by the blood collection procedure itself. Consequently measurement of cortisol concentration in saliva has become a widely accepted alternative to measurement of blood cortisol. Furthermore, the literature on salivary cortisol analysis of air traffic controllers among others, indicated that cortisol is a reliable measure of stress suitable for use in a high workload field setting (Zeier, 1994, p.13-19).

Salivary cortisol has been shown to be reliably correlated with the concentration of unbound cortisol in the blood, although the absolute concentration of cortisol in saliva is lower than in blood (Hellhammer, Kirschbaum and Belkien, 1987; Kirschbaum, Read and Hellhammer, 1992; Riad-Fahmy, Read, Walker and Griffith, 1982; Vining and McGiley, 1986). Cortisol enters saliva from blood via mechanisms that do not require active transport processes, so cortisol concentration in saliva is not a function of salivary flow rate. Furthermore, saliva is relatively free of other proteins and protein-bound molecules that may distort the neuroendocrine assays.

Many factors increase stress and influence the activity of the HPA axis and salivary cortisol concentration (Kirschbaum, Read and Hellhammer, 1994). Among the most common activities known to influence cortisol concentration is smoking. Nicotine stimulates the HPA axis and increases cortisol and ACTH during the period of smoking. Therefore, smokers often show a blunted cortisol response to stress, which may be due to chronic elevation in cortisol level and reduced responsiveness of the HPA axis. Physical exercise and physical workload also increase cortisol levels in blood and saliva. Cortisol is elevated throughout the period of increased physical workload and reaches peak levels 20 to 30 min after the physical stress has ended. In the case of very heavy physical exercise, such as a marathon race, cortisol rises continuously throughout the duration of the competition.

Circadian variation in adrenal function has long been recognized (Pincus, 1943). However, it was not until reliable specific radioimmunoassays for detecting cortisol were developed that the prominence of the circadian rhythm for this hormone was fully realized (Migeon et al., 1956; Peterson, 1957; Aschoff, 1979). In all species studied the peak of corticosteroid release occurs shortly before or at the onset of physical activity following a period of sleep and quiescence.

In humans, corticosteroid secretion exhibits a strong, or high-amplitude circadian rhythm, which means that the maximum deviation from the 24 hr mean value can be 100 percent or greater. The peak secretion occurs just before the time of awakening, then falls throughout the day with minimum values reached about midnight (Bartter, Delea and Halberg, 1962). Typically over 70 percent of the total 24 hr secretion of cortisol occurs between the hours of midnight and 0900 (Nichols and Tyler, 1967).

Interestingly, the circadian rhythm of cortisol secretion persists even when humans are deprived of nocturnal sleep. The rhythm of cortisol secretion is therefore linked to an internal circadian pacemaker, not to the end of sleep or onset of physical activity, per se (Weitzman et al., 1971). Unless the central circadian pacemaker itself is "reset" to a new phase position, e.g., by reversal of the light-dark cycle over several consecutive days, the cortisol rhythm remains linked temporally to other physiological measures such as core body temperature and volume of urine production (Czeisler et al., 1989).

3.7.2 Subjective Measure of Stress

Several methods exist to evaluate chronic occupational stress. These methods generally involve survey instruments, and the topic areas that they cover have been included in this study's background survey instrument. For this project it was desirable to document overall daily stress levels, as well as variations in stress responses over the course of the work shift. Several subjective report techniques for assessing changes in symptoms of stress among and within workdays were examined, including the Beck Depression Inventory (Beck, 1987), the UWIST Mood Adjective Checklist (Matthews, Jones and Chamberlain, 1990), the Multiple Affect Adjective Checklist (Zuckerman and Lubin, 1986), and the state scale of the State-Trait Anxiety Inventory (Spielberger, 1983).

An extensive search of the stress literature and test banks was undertaken to find a brief survey instrument that addresses the multifaceted nature of occupational stress (e.g., frustration, anxiety, feelings of depression, etc.). The search did not yield a suitable instrument. Moreover, since survey length and complexity were concerns, the study design included asking participants to simply rate their overall stress level on a scale similar to the OW scale at 2 hr intervals throughout the workday, as described in subsection 3.6.3.

3.8 Measures of Fatigue

Physiological indices, a sleep log and subjective ratings were used to assess dispatcher fatigue.

3.8.1 Physiological Measures

Two techniques for measuring physiological indices of sleep and fatigue received consideration: measuring eyeblinks with a Nightcap device and recording the sleep/wake cycle via actigraphy. The Nightcap is a small sensor that can record the frequency and duration of eyeblinks, an indicator of fatigue. It is placed on the temporal side of the volunteer's upper eyelid. Thin wires travel from the Nightcap to a data recorder that is usually placed in the volunteer's shirt pocket. The Nightcap was not selected for two reasons. First, the technology appears too unwieldy to be acceptable to participants. Second, the use of the Nightcap in studies of this sort had not as yet been validated. While the Nightcap was originally developed for research on REM (rapid eye movement) sleep, it had not as yet been used extensively in studies of fatigue.

Actigraphs, on the other hand, are commonly used in studies of fatigue. For example, actigraphs were recently utilized in a Federal Railroad Administration study of locomotive engineer fatigue (Thomas, Raslear and Kuehn, 1997). Actigraphs are approximately the size and shape of a normal wristwatch, and are continuously worn day and night. The actigraph contains a sensor to track arm movement. The gathered information can be used to determine the amount of sleep that each dispatcher is getting per 24 hr period, and when sleep periods are taking place.

Ambulatory Monitoring Inc.'s Sleep Watch was selected for the pilot study. The Sleep Watch is easily initialized, and the data it gathers is easily downloaded for review and analysis. While the data collected during the pilot study indicated that the Sleep Watch was performing accurately, the participants did not care to wear the device. Specifically, the pilot study participants felt that the actigraphs were too large, the watchband too flimsy and the device generally uncomfortable. Given this feedback, an actigraph from another vendor, Mini-Mitter, Inc., was selected for the full field study.

Minimitter's Actiwatch® satisfied the pilot study participants' complaints. It was smaller and lighter than the model used in the pilot test. In addition, it had both a better hardware-to-computer interface for field data collection and a better software package. The device was set to sample at 1 min intervals, thereby optimizing the length and frequency of data collection. The Actiwatch software analyzed the arm movement data and provided the following fatigue-related information for use in this study:

- Sleep start time.
- Sleep end time.
- Time in bed.
- Estimated sleep time.
- Sleep latency.
- Sleep efficiency.
- Number and mean length of sleep bouts.
- Movement and fragmentation index.

Appendix A contains a definition of each of these variables and the method for computing them.

3.8.2 Sleep Log

A sleep log (see Appendix B) was developed and used in conjunction with the actigraph data to assess quality of participants' sleep and to validate the actigraph data. This sleep log is based, in part, on similar instruments used by the USCG and DOT in studies of Vessel Traffic Service (VTS) watchstander and locomotive engineer fatigue, respectively. The sleep log contained items pertaining to the timing of the participant's sleep, sleep quality, the Karolinska Sleep Scale, and Naval Psychiatric Research Unit (NPRU) positive and negative mood scale. Both the Karolinska Sleep Scale and the NPRU Mood Scale were included to explore issues of acute and chronic sleep deprivation and fatigue. Participants were asked to complete the sleep log each day, shortly after awakening. It takes approximately 1 min to complete all items.

3.8.3 Subjective Measure of Fatigue at Work

To limit the number of rating items to be completed while on duty, multi-item fatigue inventories were not included in the study. Rather, as described in subsection 3.6.3, participants were only asked to rate their current level of fatigue on a scale similar to the OW every 2 hr throughout their work shift.

4. DATA COLLECTION PROCEDURES

The test instruments and data collection procedures for the field study are described in this section. They were very similar to those employed in the pilot study, which is described in Appendix C.

4.1 Use of Human Research Participants

An Institutional Review Board (IRB) was convened to conduct a human subjects review of the study design, protocol, and materials to be used in the field study. The IRB consisted of individuals with expertise in research involving human subjects. The IRB found no risk to human research participants in the proposed research design, and consequently approved the research plan for the field study.

4.2 Sites and Participants

The primary criteria for selecting field study sites were that: 1) a freight operation and a passenger operation had to be represented, and 2) it was desirable to start with a large participant candidate pool since it was expected that only a fraction of those solicited would participate in the two-week study. Since there were only a handful of dispatching centers in the United States that met these criteria, and because securing the cooperation of a railroad is a very time-consuming process, two of the three pilot study locations were selected as the field study sites because they fit the two study criteria and because railroad management and labor representatives were already willing to cooperate at each of the dispatching centers. The Mid-Atlantic dispatching center, a freight operation, and the New England dispatching center, a passenger and commuter train operation, were thus selected as field study sites.

Data was collected for two weeks (14 days) at each dispatching center. The dispatching center for the freight operation had a total of 10 desks (i.e., that dispatching center was responsible for 10 different territories), while the passenger operation was comprised of eight desks. Both railroad management (e.g., Superintendent of Operations) and local ATDD union representatives assisted in recruiting participants.

After permission to conduct the study was received from each railroad, the local union representative at each site was contacted and asked to distribute an invitational letter explaining the purpose of the study, what was expected of each participant, compensation for participation and how to volunteer to participate. A sign-up form accompanied the letter. The sign-up forms were then returned to the local union representative at the dispatching center. The sign-up form included a request for information about the volunteer's work schedule during the data collection period. This information was later used in developing a schedule for the mTAWL observation and labeling of the plastic saliva sample collection tubes.

As follow-up to the letter and sign-up sheet sent to the local union representative, each dispatcher was solicited for participation in the study. In total, about one-third of the dispatchers at each site agreed to participate: 20 dispatchers volunteered from the freight operation and 17 volunteered from the passenger operation.

4.3 Materials and Procedures

The following instruments were used in the field study:

1. A background survey containing questions pertaining to railroad dispatcher health and well-being; work scheduling and sleep habits; sources of stress in the workplace; interpersonal relationships; control over work; and quality of life.

2. Instructions for participants, including instructions for completing the sleep log, the daily workload, stress and fatigue ratings; instructions for providing saliva samples; and directions for wearing the Actiwatch.

3. An informed consent agreement that described the conditions of the study as well as the participants' rights and responsibilities.

4. Wrist-worn Actiwatches.

5. Salivary cortisol assays.

6. Daily sleep logs.

7. Subjective rating forms for workload, stress and fatigue.

8. Debriefing survey.

Based on the pilot studies, a few minor modifications were made to some of the data collection procedures. Form D and route block counts would be collected at the conclusion of the study. Collection of communications and key press information was pursued; however, after discussions with each dispatching center's computer technicians, it became apparent that this information would be too difficult to assemble. The paper-based dispatcher records provided efficient and unobtrusive access to basic workload data without further burdening participants. The specific workload variables being investigated, as well as the method of data collection, depended upon the study site due to operational differences. Both locations used paper-based Form Ds, while route block information was only available from the freight operation. Foul time permits, on the other hand, were recorded by dispatchers at the passenger dispatching center, but were not recorded at the freight dispatching center. Again, these differences are due to differences in how each railroad operates.

Information on the number of trains and other track users was collected through dispatchers' self-reports on their rating forms. Dispatchers were asked to record separately the number of trains and the number of other track users that had traversed their territory during each 2 hr block of time. The subjective workload, stress, and fatigue rating scales were not modified except for the inclusion of these two count items. As described earlier, the survey booklet used in the pilot study, which contained each day's subjective rating forms, was replaced with daily rating form packets that were picked up at the start of each shift and dropped off at the end of each shift. Daily collection also increased the likelihood of participants completing their surveys on time and without influence from previous recordings. Since some pilot study participants were surprised by the food and smoking restrictions required by the saliva collection procedure, information regarding these restrictions was explicitly provided up front when dispatchers were recruited. Lastly, an alternative wrist-worn actigraph was selected. The model chosen, a Mini-Mitter, Inc. Actiwatch AW64, was lighter, smaller, and sturdier than the model originally piloted. (See Figure 8.) It was expected that this different make and model of actigraph would be more comfortable and therefore more acceptable to participants.

Prior to the start of data collection, each participant signed an informed consent form and was given a background survey to complete. As with the pilot study, each participant was assigned a unique identification number. Participants' names were not recorded on any of the data that were collected.

Prior to the beginning of the data collection period, actigraphs were distributed, and participants were instructed on how to complete the daily subjective ratings. At this time, participants were also given a tablet of sleep logs that they were instructed to take home and fill in after each daily main sleep period during the data collection period. Lastly, the procedure for collecting the saliva samples was explained at this time. An experimenter was present at the

510-FRA-00045-2

Figure 8. Actiwatch AW64

beginning, middle, and end of the data collection periods to answer questions and monitor the progress of the data collection. Data from the Actiwatches was downloaded at each of these periods to verify that they were collecting data properly. In addition, the experimenter was available 24 hr a day via cell phone if a participant had a question or a concern that required immediate assistance. A technician was also present at the turnover of every shift to distribute the fresh saliva sample packs and to pick up and store the used samples until they could be sent to a laboratory for analysis.

Data collection took place for two weeks in late Summer at the freight operation and two weeks in early Fall 1998 at the passenger operation. The daily data collection procedure was similar to that of the pilot study. Specifically, participants completed daily baseline subjective ratings upon reporting to work. They then provided ratings at 2 hr intervals until the end of the shift (i.e., 2 hr into shift, 4 hr, 6 hr, and 8 hr). Saliva samples were taken at the same intervals. In addition to the subjective ratings, participants completed a sleep log each day, even if they did not report to work, and wore the actigraphs at all times unless they were showering/bathing/swimming. Finally, two trained observers collected mTAWL data on one dispatcher per shift per day. To maintain vigilance, observers alternated their observation every other hour. Ten shifts were observed using the mTAWL procedure at the freight operation site, and 12 shift periods of data were collected at the passenger operation. At the conclusion of the study, participants completed a debriefing survey, returned their actigraphs to the experimenter, and were paid and thanked for their participation.

Data from one shift out of 42 total shifts (14 days x 3 shifts per day) at the freight operation were compromised due to a temporary problem on the rail network. No other data-related problems were encountered during the two weeks of data collection at either site.

Appendix B contains copies of all of the forms used in the field study.

5. DISPATCHER ASSESSMENT

This chapter presents study findings and observations about the sampled railroad dispatchers. The results are organized under five subtopic headings:

- Dispatcher Characteristics.
- Workplace Characteristics.
- Workload.
- Stress.
- Fatigue.

These data were collected from a small, non-randomly selected sample of dispatchers, and may not, consequently, be representative of the entire railroad industry. Thus, this information should be used to: 1) gain insight into the dispatchers who participated in this study, and 2) identify areas that could benefit from further research.

The majority of the significance testing was accomplished through Analysis of Variance (ANOVA) and *t*-test procedures. As this is an exploratory analysis, specific a priori hypotheses were not postulated, and statistical testing was two-tailed unless otherwise noted. A significance level of $p<0.05$ was set for this study except where otherwise stated. SPSS 8.0 (SPSS, 1998) was used to analyze the data.

5.1 Dispatcher Characteristics

This section provides demographic information about railroad dispatchers based on self-report data collected from responses to the background survey. Data from the passenger and freight operations field sites are compared. Data are also collapsed and parsed into moderating variables (e.g., age, job tenure), and compared to national normative data when appropriate.

While there is no reason to hypothesize differences between the two field locations, differences from the general U.S. population were anticipated for those variables identified by the literature as being sensitive to shiftwork. These variables include issues of psychological well-being, certain health-related afflictions such as gastrointestinal distress, headaches, and sleep problems, and a general tendency to be overweight due to poor nutritional habits when working nights. It was also expected that some health and well-being variables would be connected to job tenure and age. Job tenure was considered a relevant moderator variable since people unsuited for railroad dispatching, or more broadly, shiftwork in general, are expected to self-select out of the profession over time. Age was considered a moderating variable, since age-related physiological changes affect susceptibility to various illnesses and conditions.

5.1.1 General Background Information

Table 5 summarizes several dispatcher demographic variables. Analysis of Variance (ANOVA) revealed no statistical differences between the freight and passenger operations on these variables. All of the study participants had at least a high school diploma, and the majority had at least several years of college (70 percent at freight, 82.4 percent passenger dispatch operations). Out of the 37 total participants, 31 were male (84 percent), with two female dispatchers working at the freight operation and four at the passenger operation. Dispatcher participants at the freight and passenger operations weighed, on average, 142.5 and 192 lb, respectively. Sixty-five percent of the freight dispatchers and 63 percent of the passenger dispatchers were married, having, on average, 1.5 and 1.6 children, respectively. The median ages of dispatchers at the freight and passenger operations were 43.5 and 43 years, respectively. Participants were also asked how old they feel, a measure known as subjective age. Research has found that a higher subjective age (relative to physiological age) is related to chronic stress and poor psychological well-being (Barnes-Farrell and Petrowski, 1991). Given the nature of the job, it was expected that participants would report greater subjective ages relative to their true ages. In fact, a significant difference was found between how old participants were and how old they felt. However, at both sites, participants' subjective ages were generally lower than their physiological ages. Looking at the data separately for those younger than 40 and those older than 40 produces similar results.

To summarize, the typical dispatcher in this study was experienced in their job, male, married with 1 to 2 children, had at least a high school diploma, and was overweight[2]. This latter finding is particularly noteworthy, since overweight people tend to incur increased sleep

Table 5. Demographic characteristics of the study participants

Variable	Overall (n=37)
Median Age (Median Subjective Age)	43 (35)
Sex (percent male)	84%
Weight	184
Median Job Tenure (months)	100
Median Months as Dispatcher	96
Percent Moonlighting	8%
Percent Married	70.3
Mean Number of Children	1.5
Education (range in years)	12 to 16 years
Percent Consuming Caffeine	89%
Mean Caffeine Usage (cups/can per day)	3.2 cups/cans
Percent Smokers	28%

[2]This was determined using the 1999 Metropolitan Life Insurance Company height weight charts for men and women found at http://www.metlife.com/Lifeadvice/Tools/Heightnweight/index.html

difficulties such as sleep apnea, which in turn influence their wake-time alertness and sleepiness levels. A second interesting finding was that participants reported feeling younger than they were, suggesting that the dispatchers who participated in the study may not feel chronically stressed. The issue of stress is discussed in detail in subsection 5.4.

5.1.2 Job Experience

The median[3] tenure for the participants' current job was about 100 months. Twenty-three of the 37 total participants (62 percent) have five or more years of experience, and are generally considered to be "veterans" by the industry. The percentage of experienced dispatchers who participated in the study was similar at both sites. The experience of the remaining 14 participants ranged from just beginning their career to having 58 months of experience. Only three of the 37 dispatchers, all from the freight operation, reported working a second job. These second jobs were home or organization-based (e.g., embroidery work, church organist) and do not appear to demand large amounts of the dispatcher's time. Approximately four out of every five (81 percent) participants worked in another capacity for the railroad before entering dispatch work. Table 6 provides a breakdown of these different jobs. The non-railroad prior work experiences varied widely and included such occupations as student, meat packing and restaurant management. Those dispatchers with less job tenure tended to be the ones with less prior railroad experience before becoming a dispatcher.

Examination of job experience-related variables yielded two interesting findings. First, very few of the participants worked second jobs. The lack of second jobs may be the result of the availability of abundant overtime, the unpredictability of changing work schedules, or simply that they are satisfied with the money they currently earn through dispatching. Another reason may be participants' desire to spend time with their families. The second interesting finding was that many of the less experienced dispatchers came from jobs outside the railroad industry. This provides evidence supporting one of the concerns recently expressed by the railroad industry, which is that the traditional career path to becoming a dispatcher is being eliminated. Previous entry points to becoming a dispatcher, such as working as tower and block operators, are being eliminated due to advances in dispatching technology and operational changes such as centralized dispatching. As a result, the railroads are beginning to have to hire people "off-the-street" for one of its most demanding, safety-critical positions. These new hires do not have any prerequisite knowledge of railroad operations. Only seven dispatchers in the current study fit this description. This sub-sample was considered to be too small, both with regard to the current sample as well as with regard to the U.S. dispatcher population,

Table 6. Prior work experience of study participants

Prior Positions	Number (%)
Block Operators	11 (29.7)
Railroad Clerk	10 (27.1)
Railroad Operations	6 (16.2)
Other Railroad Positions	3 (8.1)
Non-railroad	7 (18.9)

[3]The median was used instead of the mean as a measure of central tendency. This was done as an attempt to mitigate skewing that might occur should either a preponderance of younger or older dispatchers confound the study of interest.

approximately 2500[4], to provide meaningful results, and thus is not considered further in this report.

5.1.3 Dispatcher Health and Well-Being

The prevalence of illnesses and conditions associated with stress and fatigue within the dispatcher sample was evaluated next. Though it is risky to extrapolate these results to the entire dispatcher workforce, examination of health and well-being issues among study participants is useful to determine if there are any outstanding issues that might indicate the need for further examination.

First, comparisons between U.S. population norms and participant responses to a number of health-related questions were made. U.S. population data came from the Centers for Disease Control's National Center for Health Statistics' "Fastats" (http://www.cdc.gov/nchs/fastats/fastats.htm, June 1999) and are reported as rates per thousand adult Americans. Study participant responses were collapsed across the two field sites and converted into incident rates per thousand for comparison purposes. When U.S. data was available by age group, separate incident rates were calculated for those 25 to 44 and those 45 to 59. Table 7 presents incident rates per thousand individuals for both the study population and U.S. population norms (in parentheses). An epidemiological statistical method was then applied to the data to determine if there were statistically significant differences between the study population rates and U.S. norms. For each

Table 7. Comparison of dispatcher health, incidence per 1000, to U.S. norms

	Age Range	
	25-44 (n=26)	45-59 (n=11)
Anxiety	115.4 (126)	90.9 (126)
Asthma	0 (60.7)	0 (31.4)
Back Pain	307.7 (42.4)*	272.7 (110.3)
Depression	38.5 (50)	181.8 (50)
Gastrointestinal	38.5 (2.7 to 24.9)	90.0 (7.2 to 33.5)
Headaches	230.8 (21.8)*	90.9 (31.7)
Heart Disease	38.5 (24)	0 (143.1)
Hypertension	38.5 (34)	181.8 (233.2)
Skin Disorders	153.8 (5.4 to 33.3)*	0 (4.1 to 26.8)
Sleep Problems	269.2 (350)	363.6 (350)

( ) Indicates U.S. population norms.
* Indicates $p<0.05$ significance level.

[4]This value is based on counts provided by the major railroads and dispatcher unions.

health-related variable, a critical incident rate was calculated (see Equation 2), based on the incident rates for both the study population and the U.S. norm. This critical incident rate was then compared to the critical value for a normal two-tailed distribution ($p < 0.05$), 1.96. Statistically significant differences occurred when the critical incident rate exceeded 1.96. The results of these analyses are presented in Table 7. An asterisk in Table 7 indicates a significant difference between the study population incident rate and U.S. norms for that particular health-related variable.

$$\text{Critical rate} = \frac{r_s - r_p}{\text{StD}} \tag{2}$$

Where:

r_s = rate of sample
r_p = rate of population and

$$\text{StD} = \sqrt{\frac{r_s \times (1 - r_s)}{n}}$$

As can be seen, the younger study participants reported experiencing significantly higher rates of back pain, head aches and skin disorders compared to the general U.S. population. While causation cannot be determined from these data, it seems that these ailments are as likely due to the physical environment of the workplace as due to job-related stress. Interestingly, the older study participants did not report significantly higher incident rates for any of the health-related variables. Although it is possible that older participants did not have any significantly worse health-related problems compared to U.S. norms, it is also possible that the lack of significant differences may be a function of low power due to such a small sample size (n=11). Regardless, there appear to be some problems with the younger dispatchers in the study which may merit further investigation into health-related problems within the entire dispatching population. Only three conditions, anxiety, depression and sleep problems, were able to be collapsed over age groups due to similar normative values. This larger sample size increased the power of the analysis, revealing significant critical values for anxiety (critical value = -2.36) and sleep problems (critical value = -4.36).

5.1.4 Quality of Life

Table 8 and Table 9 summarize participants' responses to 14 different aspects of their lives. The tables are separated into positive and negative feelings. Nearly all participants, 84 percent, felt confident handling personal issues "fairly often" or "very often." About half, 51 percent, felt on top of everything "fairly often" or "very often," and only 41 percent felt they were in control of the usage of their time "fairly often" or "very often." These two results may be due to the need or pressure to work overtime. Furthermore, it appears that participants are bringing home workplace stress. Only half of the participants reported feeling that life is going well "fairly

Table 8. Quality of life: positive factors

Feeling	% Fairly or Very Often
Confident Handling Personal Issues	84
Able to Cope	58
On Top of Everything	51
Positively Dealt with Hassles	49
Life Going Well	49
Able to Control Life's Irritations	46
Able to Control Time Usage	41

Table 9. Quality of life: negative factors

Feeling	% Fairly or Very Often
Preoccupied with Things to Do	51
Angry about Uncontrollable Issues	35
Stressed	27
Unable to Control Life	19
Upset over Something Unexpected	16
Unable to Cope with Responsibilities	14
Overwhelmed with Difficulties	11

often" or "very often," and slightly less than that felt they were able to control life irritations "fairly often" or "very often."

Table 9 suggests possible negative effects of overtime and workplace stress. Half of the participants reported feeling preoccupied with things to do "fairly often" or "very often." This may be due to working more than 40 hr per week, which results in less free time for personal activities. Furthermore, as there is no break period and the work is done in an open environment, it seems unlikely that the dispatchers are able to take care of much personal business while at work. About one-third of the participants reported that "fairly often" or "very often" they had been angered by uncontrollable issues and a quarter reported feeling stressed fairly or very often. Unfortunately, national norms were not available with which to compare these findings.

5.1.5 Summary

As expected, with respect to dispatcher characteristics, there were no substantial differences between the two field study locations. This was a fairly homogenous group, and likely reflects the largely homogenous railroad work force. The railroad industry is one of the oldest, and its workforce has been made up largely of married, white men with high school degrees. There are signs, however, that the makeup of the workforce is changing, especially as new dispatchers (among other railroad crafts) are hired off the street rather than through an established career path. The effects on dispatcher performance and work stress due to this change in career path are currently unknown. Among other adjustments that railroads have made or will have to make in the near future, dispatcher training has had to change to meet the needs of new hires who lack the prerequisite knowledge of railroad operations that previous generations of dispatchers had when beginning their dispatching careers.

The dispatchers who participated in this study, especially the younger ones, did show some characteristics similar to other shiftworkers (i.e., sleep problems, headaches, trouble managing time), in addition to some other problems (i.e., anxiety). Specifically, it appears that the dispatchers work a lot of overtime. This is potentially problematic in that overtime precludes the dispatcher benefiting from the recovery time built into the functional work schedule system. Lack of recovery time can lead to feelings of stress, anger towards the workplace, and a sense of never having enough time. Such findings were present in this sample. Furthermore, a very demanding, high stress work environment could explain the health and stress-related findings. The exact etiology of the stress, whether it comes from work pace, interactions with rail personnel, the physical environment, or other sources, is initially less important than understanding whether the work experience, in fact, is producing a significant, chronic stressor reaction that is impacting other aspects of the workers' lives. This issue will be more fully addressed in the following sections. The data presented here, however, does indicate evidence of high stress in a significant number of the study participants.

5.2 Workplace Characteristics

The data presented in this section are also based on dispatchers' responses to the background survey. Passenger and freight operations are compared, and data are collapsed and grouped by moderating variables (e.g., age, job tenure) when appropriate.

Based on anecdotal information from dispatchers, dispatchers have the opportunity to work a lot of overtime. In addition to the overtime, dispatchers' daily commute time impinges on the amount of free time that a dispatcher has over the course of a week. The result is that a dispatcher may have less than an ideal amount of free time. Further, while official rest breaks are not built into most railroad dispatchers' work, it is likely that dispatchers do, in fact, take breaks, though the number of breaks, and the exact reasons for taking them, are not known. Lastly, the safety critical nature of the work, and the physical and social environment of the workplace, are expected to contribute to dispatchers' feelings of stress.

5.2.1 Work Hours

Most dispatchers (89 percent) reported that, on average, they are scheduled to work a consecutive five-day workweek. However, over half of the dispatchers also reported that they work an average of eight or more hours of overtime weekly. In fact, two participants had worked more than 30 consecutive days prior to the start of the study. Twenty-seven percent of the participants felt that they were expected to work overtime. This figure rose to 42 percent when considering just those dispatchers with less than two years of job tenure. These less-tenured dispatchers tended to work primarily the extra board and were paid only 75 to 80 percent of a full dispatcher salary. Thus, not only do these less-tenured dispatchers have less control over their work hours, their lower salaries may serve to motivate them to work extra hours.

The high rate of overtime may be due to understaffing at the two study sites. However, this issue was not examined in the present study, and thus it is not possible to determine the possible contribution of staffing levels to the high rate of overtime experienced at these two study sites. Regardless of the cause(s), excessive overtime often causes workers to obtain less recovery time away from work than what is needed. Days off are an integral part of the work schedule. They are designed to provide the worker with time to recover physically, emotionally, and cognitively from their work. Time off also provides workers with time to address the other areas and needs in their lives. Ignoring this recovery component, especially in a chronic fashion, has the potential to increase the chances of work-related fatigue and burnout, and family problems.

5.2.2 Commuting

Like overtime, relatively long commute times also consumed the dispatchers' off-time. Figure 9 presents time at work and commute time for the two field study sites. On average, the dispatchers at the passenger operation reported being at the job site for 8 hr and 27 min while the freight operation dispatchers reported being on site for 8 hr 13 min per shift. An independent *t*-test did not reveal a significant difference between the two sites. There was a statistically significant difference in commute time, however. Specifically, the average round trip commute for the passenger operation dispatchers was 2 hr 12 min, almost three times the length of the average freight operation dispatcher commute time, which was 50 min (t=-2.841, df=35, p=0.007). Adding time at work and commuting time together reveals that, on average, the passenger dispatcher is committed for 10-1/2 hr during a workday, while the freight dispatchers are committed for about 9 hr. This difference is statistically significant (t=-2.747, df=30, p=0.01).

The above difference is likely to be not a function of type of operation, but rather, one of dispatching center location. The passenger operation was located in a major metropolitan area, while the freight operation was located in a more rural setting. If 7 hr of sleep per night is assumed, then the passenger operation dispatchers have about 5 to 6 hr of free time per workday

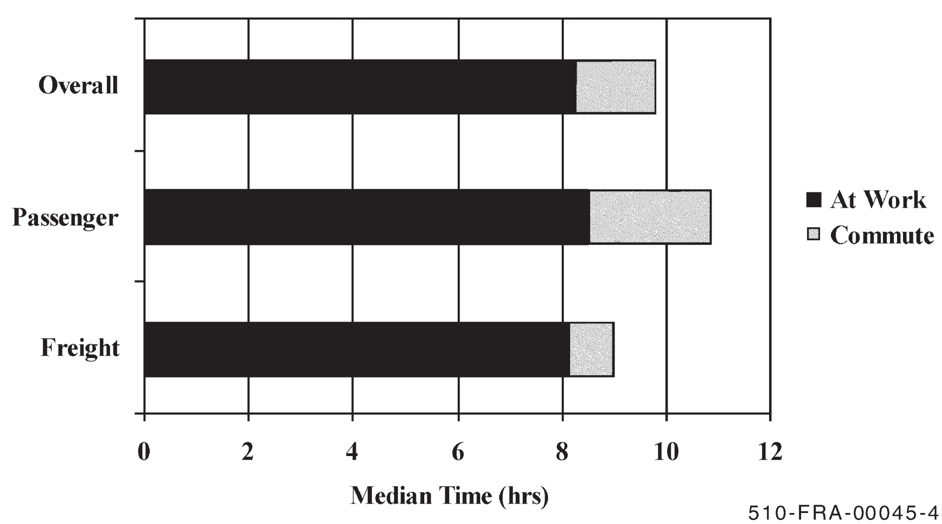


Figure 9. Work-related daily time

to eat, be with family and friends, and take care of personal business. From this is it easy to see how inadequate recovery time might occur if the dispatchers are also required to work at least one of their scheduled days off. As mentioned above, chronic exposure to this type of situation might lead to burnout, stress, and a host of family and other personal problems.

5.2.3 Rest Breaks

Although there are no official rest breaks, participants reported that they did, in fact, take breaks from their desks. The median[5] number of breaks taken per shift was four, ranging from 0 to 10 for both study sites. The median number of rest breaks for smokers, who comprised 12 of the 37 participants, was 3.5. Ninety-seven percent of the dispatchers reported eating at their desk, while the other 3 percent of participants reported that they did not eat at work. It appears, therefore, that rest breaks were taken for a more urgent reason, such as toileting or the "need to break away and clear the mind."

Even though it seems that rest breaks are taken, they are not necessarily taken when desired. Forty-one percent of the dispatchers reported that they are “frequently” too busy to take a break while 57 percent felt that at least “occasionally” they were too busy to take a break when they wanted. As a further complication, 65 percent of the respondents reported that they had “frequent” or “constant” difficulty finding someone to cover their desk in their absence.

These data suggest that it is difficult for a dispatcher to take a break, and that these breaks often do not occur when they are most needed or wanted. Further, when breaks are finally taken, it is often difficult to find someone to cover the desk during that time. An institutionalized system for taking rest breaks, something which did not exist at either field site, might help alleviate the problem of taking urgent breaks (i.e., toileting) without compromising safety. The larger issue of whether institutionalized breaks, including one for lunch, would improve performance over the course of a shift, remains unanswered.

5.2.4 Job Demands

Table 10 presents dispatchers’ ratings on their perceived volume and pace of work. Of note is that 97 percent of the dispatchers reported that they were “often” or “very often” confronted

Table 10. Volume and pace of work by shift

	Percent Reporting Experiencing “Often” or “Very Often”				
	Day	Evening	Night	Extra Board	Overall
Working Fast	75	83	80	91	84
Large Voume of Work	92	67	100	64	81
Heavy Concentration	92	100	100	91	95
Heavy Memory Usage	92	100	100	100	97
Lulls between Heavy Periods	25	17	20	27	22

[5] The median was used instead of the mean as a measure of central tendency. This was done as an attempt to mitigate skewing that might occur given the large variance in number of breaks taken over the course of the work shift.

with heavy memory usage, supporting the previous finding that 98 percent of the dispatchers reported being “occasionally” or “frequently” too busy to break. Neither of the dispatching centers in this study used computer-aided dispatching systems that provided memory or decision aids.

Despite this reported high workload, 95 percent of the dispatchers reported feeling competent to handle emergency situations. Furthermore, 86 percent of the dispatchers reported that they “often” or “very often” knew exactly what was expected from them. Predictably, the 14 percent of respondents who reported that they “rarely” or “occasionally” knew exactly what was expected from them were those with less work experience. Separately, only 14 percent of participants felt that they were “often” or “very often” given assignments without the support they needed; these dispatchers tended to have greater job tenure. Nine percent of those surveyed reported “often” or “very often” bending rules to be able to complete the job; all of these respondents had five or more years of job experience. Only 11 percent responded that they were “frequently” or “always” asked to perform tasks outside of their job guidelines.

5.2.5 Workplace Dynamics and Stress

The issue of stress in the workplace was examined from four perspectives: emotional demands, physical work environment, perceptions of workplace control, and interpersonal interaction. Each is discussed below.

5.2.5.1 Emotional Demands

Nearly all the dispatchers, 97 percent, reported feeling a great deal of responsibility for the lives and welfare of others, though only 43 percent claim that this caused “a lot” or a “great deal” of stress or anxiety. Still, 30 percent reported “occasionally” and 9 percent reported “frequently” calling in sick due to stress. While almost 40 percent of the respondents reported that they call in sick at least occasionally due to stress, national norms were not available for comparison.

5.2.5.2 Physical Work Environment

Table 11 reveals that there was some discomfort with regard to the physical work environment. This was expected, though, given the variability in human comfort levels. Of note is the rating for air quality; less than 50 percent of the dispatchers reported air quality as acceptable. Both dispatching centers were in windowless office structures and therefore relied on internal ventilation. All respondents reported that they can and do adjust their chairs when coming on shift. Furthermore, all dispatchers reported a “high” or “very high” visual demand component to their work, monitoring between four to seven electronic visual displays throughout their shift. Half of the dispatchers reported feeling at least “sometimes” having irritated, itchy, or sore eyes.

Table 11. Physical environment

Factor	% Acceptable
Lighting	70
Noise	62
Temperature	57
Air Quality	49

5.2.5.3 Perceptions of Workplace Control

Dispatchers reported having the most control over work quality and task ordering, and the least control over policy decisions. See Table 12 for details. There is literature suggesting an inverse relationship between a worker's perception of control over his or her work and feelings of stress (Selye, 1978). Lower perceptions of control are associated with higher stress levels.

In general it appears that given the nature of the job, and its tight government and company regulation, that these dispatchers do not feel that they exert a high level of control. Job stress frequently results when job-decision latitude is not commensurate with the psychological demands of the job (Sauter and Murphy, 1995). This potential stressor, low control over one's work, is not likely to change in the near future. Nevertheless, educating the dispatchers as to the reasoning behind regulations and policies may alleviate some of the negative, stressful feelings attributed to lack of control over policy decisions.

5.2.5.4 Interpersonal Interaction

On the positive side, these dispatchers seemed to have a strong support network. Over half reported the chief dispatcher (62 percent) and coworkers (57 percent) made life easier at the center. During hard times caused by work, these dispatchers felt that they could rely most on their coworkers (78 percent), then friends and relatives (73 percent) and the chief dispatcher (68 percent). As to be expected, 95 percent of the dispatchers reported an ease of talking out problems with their friends and relatives, followed by 91 percent for coworkers and 84 percent for the chief dispatcher. A former railroad dispatcher expressed the importance of these associations when interviewed about personal experiences. This dispatcher felt that it was imperative for new dispatchers to be able to fit into the social environment present in the office. Compatibility with the existing social structure at the center enables the dispatcher to successfully handle the rigors of the job itself without quickly burning out and leaving the position.

Table 13 presents the perceived level of conflict and cooperation between the dispatchers and those with whom they interact at work. It is interesting to note train crews are reported to cause the most conflict for the dispatchers, but were also the most cooperative. This is logical as both these groups realize that working together is in their mutual interest to keep the trains moving safely, yet both are under pressure to accomplish their job as quickly and efficiently as possible.

Table 12. Workplace control issues

Factor	% Reporting Moderate or High Level of Control
Work Quality	97
Task Ordering	81
Number Simultaneous Tasks	57
Work Pace	43
Policy Decisions	16

Table 13. Dispatcher perceptions of levels of conflict and cooperation with other railroad workers

Dispatcher Versus	Level of Conflict (%)	Level of Cooperation (%)
Train Crews	49	95
Maintenance Crews	43	84
Operations Manager	41	81
Other Dispatchers	38	87
Chief Dispatcher	24	89

The data indicated that the operations manager was associated with the most problems for these dispatchers, with a moderate level of conflict and the lowest level of cooperation among the comparison groups. It is likely that the operations manager is the least involved in making sure the daily work is accomplished and is likely to intervene only when there is already a problem or a perceived situation. The other four groups have to work together more often, making sure the work is accomplished on a day-to-day basis.

Despite the apparent strong social support network and cooperation, 27 percent of the dispatchers did report "sometimes" or "frequently" losing their temper at work. In addition, 92 percent reported that others on their shift "sometimes" or "frequently" lost their temper. The managers of the study sites provided anecdotal information that supported this result. Figure 10 shows that over 40 percent of the participants report that they take work stress home with them. While this percentage was higher for the passenger operation dispatchers, this difference was not statistically significant.

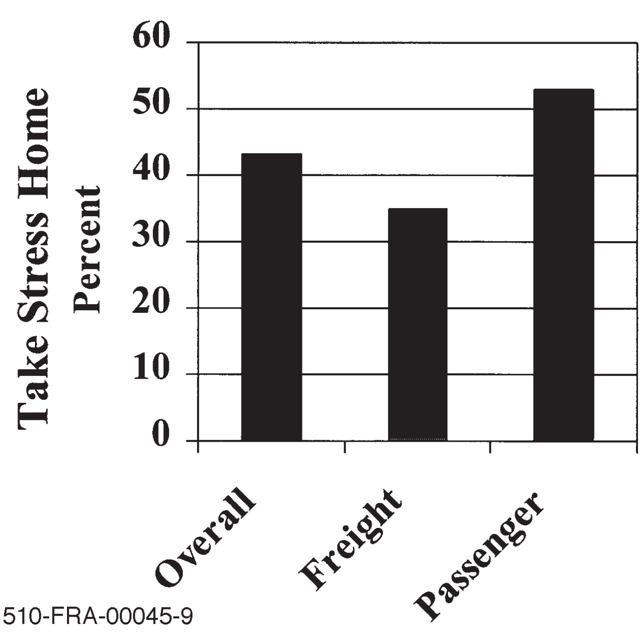


Figure 10. Evidence of workplace stress

5.2.6 Summary

Most full-time occupations have work schedules with built-in lunch or rest breaks. While the dispatchers' work shifts are typically shorter than those in other full-time occupations due to the lack of planned breaks, the commute times tend to average out this difference, especially with regard to the urban field site. Also, as there is so much overtime being worked, the dispatchers' amount of free or recovery time is further diminished, possibly to a point of breakdown in the recovery process. Such a breakdown might manifest itself in the form of increased stress and other problems both inside and outside of work. This issue is further compounded by the lack of formal breaks while at work, and that the breaks that are taken are usually not when desired. Given the difficulty in finding someone to cover a desk, these breaks may be taken without anyone monitoring the territory.

Regarding the job and work environment, though nearly all the dispatchers reported that they often dealt with high workloads, most felt capable of handling emergencies and competent in handling all aspects of their job. Only a few felt it necessary to bend the rules to accomplish their work or perform work outside their job guidelines. Though less than half felt that the nature of the job itself causes much anxiety within them, a large minority of the participants did report taking sick time due to stress. The nature of this stress seemed to result less from the physical work environment and more with issues of control over work and interpersonal interactions. This is a somewhat different finding than what was expected, that workload and work responsibility would cause the majority of the work-related stress.

5.3 Workload

Dispatcher workload was examined using three separate sources of data. The first source of data, considered to be the most objective, was derived from the mTAWL. These data were collected through naturalistic observation of the dispatchers performing their job. The second type of workload data was related to traffic volume and was obtained from dispatcher self reports of the number of trains and other track users they handled every 2 hr over the course of their shift. Finally, subjective rating data were also collected from the dispatchers every 2 hr during their shifts. These subjective ratings included two NASA TLX subscales—mental demand and level of effort—and an overall workload scale.

Again, the paucity of data currently available on railroad dispatcher workload precluded any specific a priori hypothesis to be made. Based on anecdotal information, however, it is likely that there would be time-of-day, territory, and type-of-operation (e.g., passenger or freight) differences. For example, it was expected that the passenger operation would show rush hour effects during the workweek, an event that probably does not occur on the territory controlled by the freight operation. Furthermore, it is likely that not all desks are affected by rush hour at the passenger operation, or at least at the same time or intensity. There is no available, published evidence to date, however, that reveals workload levels, even during rush hour, as being excessive and beyond safe human handling.

5.3.1 Objective Workload

The mTAWL methodology is the most objective measure of workload employed in this study. As mentioned, trained observers recorded all activities and track user movements that occurred during a dispatcher's shift. Given the time-intensive nature of this methodology, and being that this was the first time it was employed in this manner, the data it provided was used primarily in a diagnostic manner. That is, the mTAWL was used to evaluate relative workload levels of the dispatcher desks at the two field sites. It was not used to collect information for the purpose of dispatcher comparison or correlation to dispatcher subjective reports.

As discussed under Section 3, the first step of the mTAWL analysis was to hold a small focus group with experienced dispatchers at each dispatching center. The purpose of this focus group was to obtain consensus on the different dispatching tasks and on the auditory, visual, cognitive and psychomotor channel loadings to be used in the analysis. Table 14 presents the various mTAWL category channel weights for the two study sites. The two field sites independently derived channel loadings for all presented categories.

Appendix C contains mTAWL charts for each participant that was observed during the data collection period. The 22 mTAWL charts present standardized data for a participant relative to the other desks that were observed at his or her dispatching center. Given that most desks were evaluated only once, on a randomly selected shift and day, and thereby leaving incomplete cells in the data matrix, these data were not aggregated for statistical analysis. Rather, these data were used for diagnostic purposes to determine workload trends within each dispatching office.

5.3.1.1 Shift Effects

Visual inspection of these data (see Appendix D) revealed a possible desk by time-of-day by time-of-week interaction. For example, Desk A of the freight operation on the Day 2 day shift, a Tuesday, shows a different workload pattern than the Day 5 day shift of this same desk (see Figure 11 and Figure 12, respectively). In this case the work surge seems to come earlier on Day 5 than on Day 2. Furthermore, if Desk A day shift is compared with data from its night shift recorded on Day 12 seen in Figure 13, the change in overall level of activity is immediately noticeable. There also seem to be differences between desks. While all the desks observed on the first shift seem to show workload peaks, the desks observed on the second shift show considerable variation. For example, Desk D hovered around the mean for the duration of the shift while desk G vacillated, peaking twice for short periods of time, while Desk F peaked once during the first third of the shift, while Desk E showed high workload throughout the shift (see Figure 14 through Figure 17). The desks observed during the third shift showed similar differences, though the absolute level of activity seemed to be less. Interestingly, there appear to be distinct channel groups, that is, the visual and auditory channels and the cognitive and psychomotor channels tend to cluster together at about the same amplitudes. These two groupings, however, tend to have different amplitudes. It is unclear at this point, however, if these differences are due to time-of-day, desk demands, job experience, or some other factor.

Table 14. mTAWL channel loadings

		Channels			
		Auditory	Visual	Cognitive	Motor
Background Auditory Monitoring	F	1	0	0	0
	P	2	0	0	0
Background Visual Moitoring	F	0	1	1	0
	P	0	2	0	0
Background Radio/Telephone	F	3	1	2	2
	P	4	1	2	1
Foreground Radio/Telephone	F	3	1	2	2
	P	3	1	3	1
Background Clerical	F	0	2	2	6
	P	0	2	2	3
Planning for Unopposed Track Movement	F	1	1	1	1
	P	1	1	2	0
Handling Unopposed Track Movement	F	1	1	1	1
	P	1	2	2	2
Planning for Opposed Track Movement	F	1	1	3	1
	P	2	1	3	0
Handling Opposed Track Movement	F	1	1	3	2
	P	3	1	3	2
Foreground Clerical	F	0	2	2	6
	P	2	2	3	5

Key: F = freight
P = passenger

There seems to be a bit more within shift variability between desks for the passenger operation dispatchers. For example, the two times Desk A was observed on the day shift, levels of workload were primarily around or below the mean. Desk B, however, showed a late morning peak while Desk C began an extended interval of extremely high workload soon after the dispatcher came on shift. There are also between-shift differences, as observed by comparing the charts of Desks B and D between first and second shift. Second shift and third shift activity observed on Desk G appear to be much more quiet compared to first, and show a much more defined period of activity than the data charted from the freight operation. Of note is that this set of charts does not indicate a grouping of workload channels as seen with the freight operation.

5.3.1.2 Between-Desk Comparisons

While these charts provide the pattern and amplitude of workload over the course of a shift, they do not provide an easy way to directly compare desks at a given center. Table 15 extracts

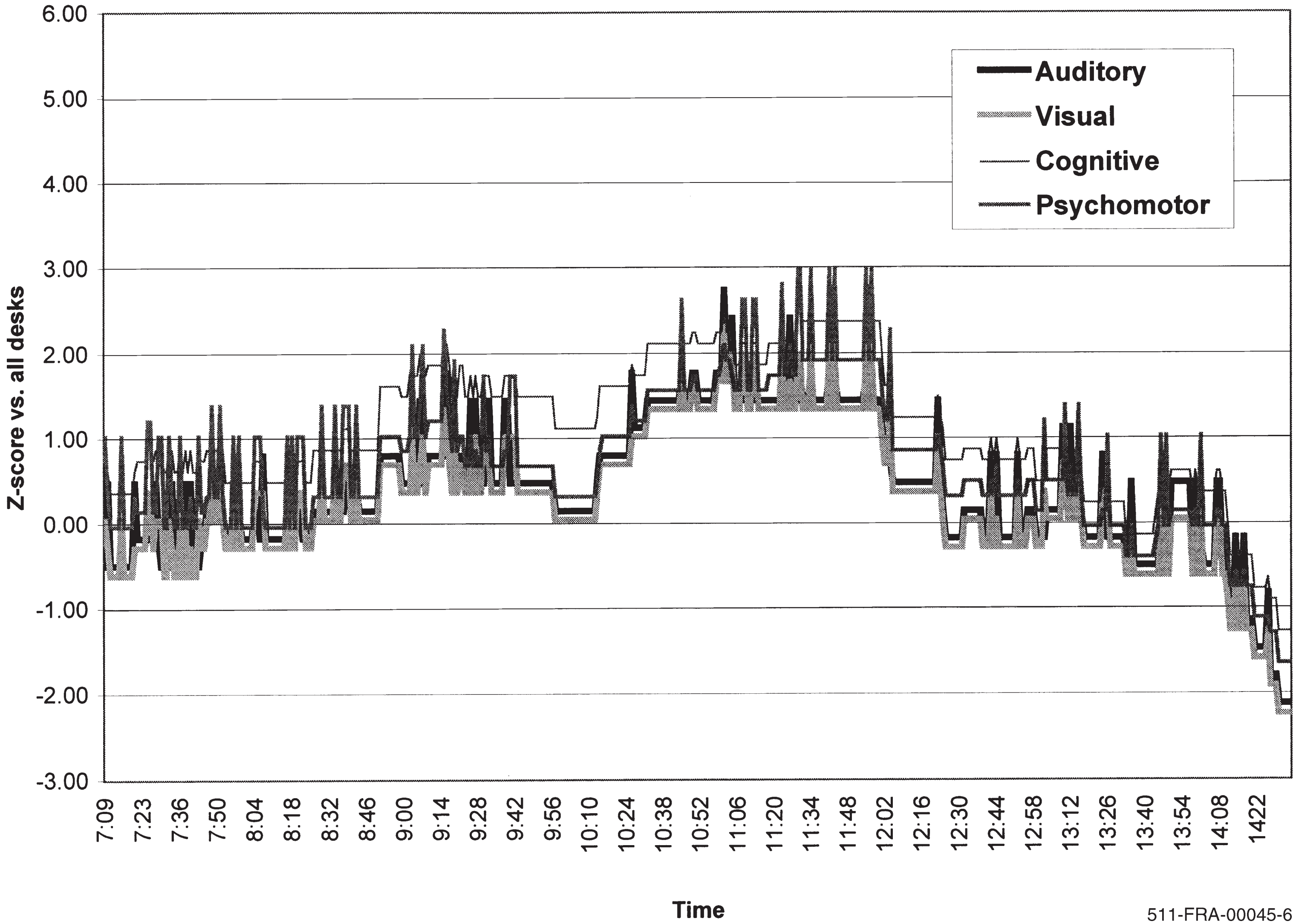


Figure 11.   Freight operations, Tuesday, day shift, Desk A

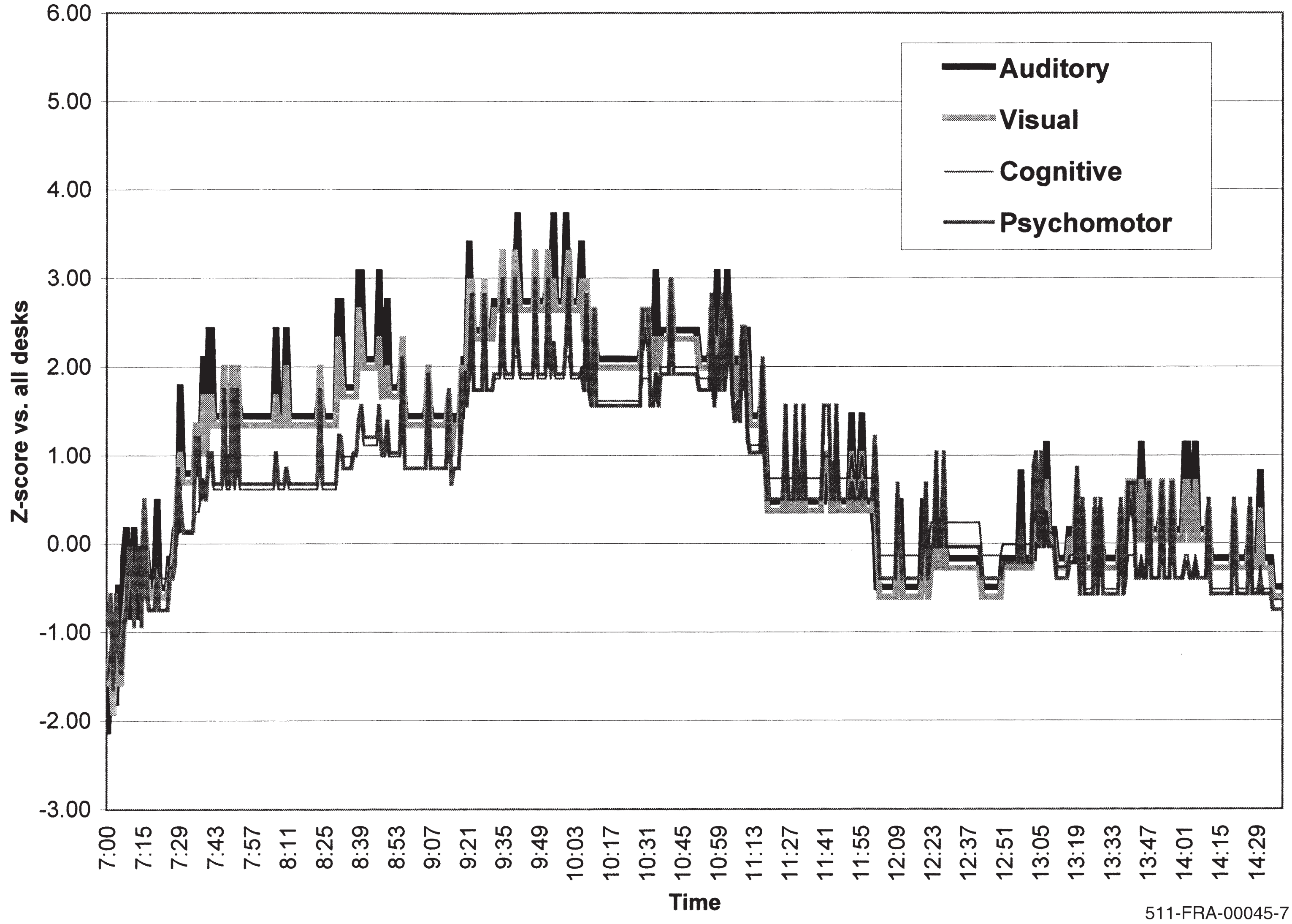


Figure 12. Freight operations, Friday, day shift, Desk A

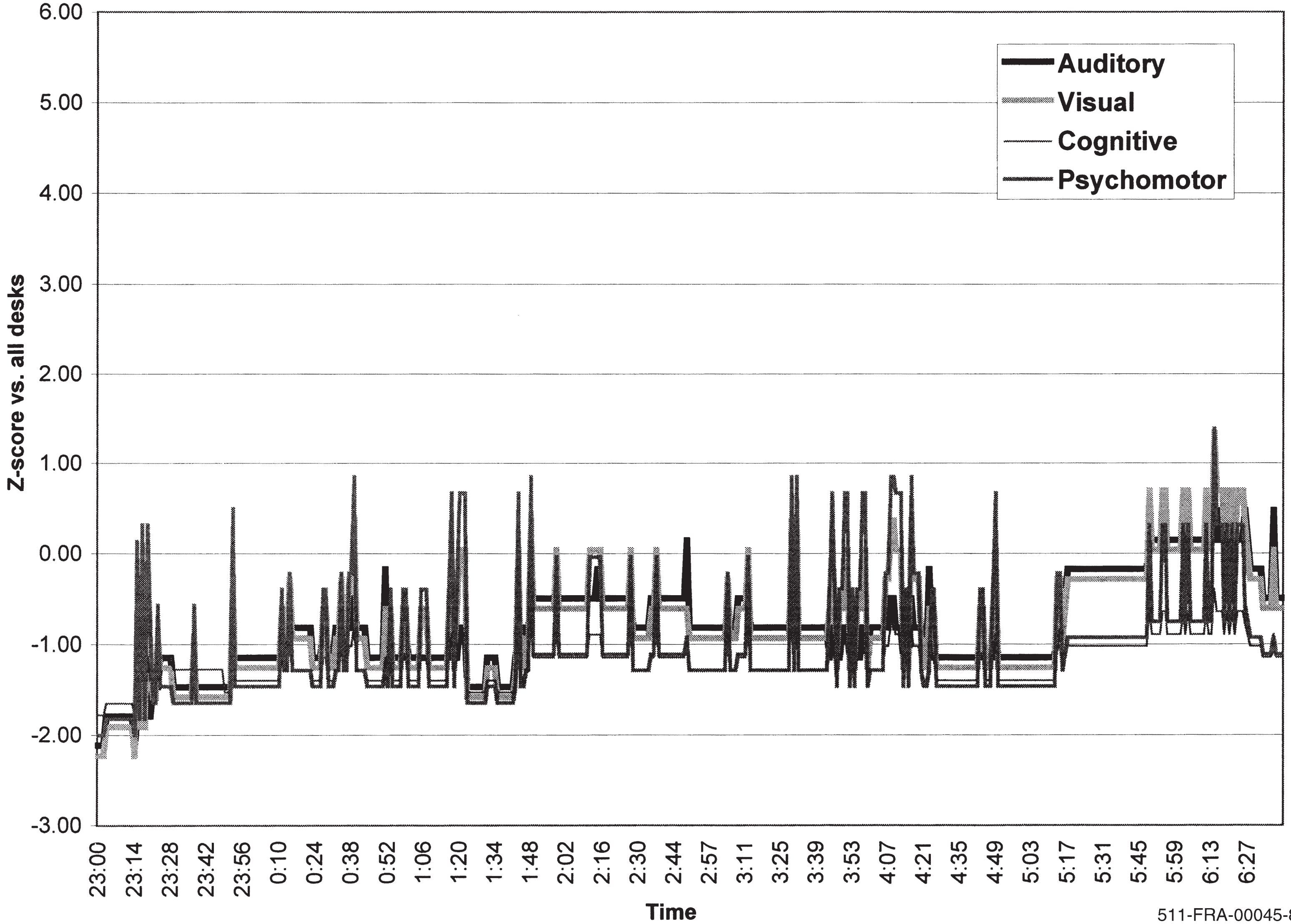


Figure 13. Freight operations, Thursday, night shift, Desk A

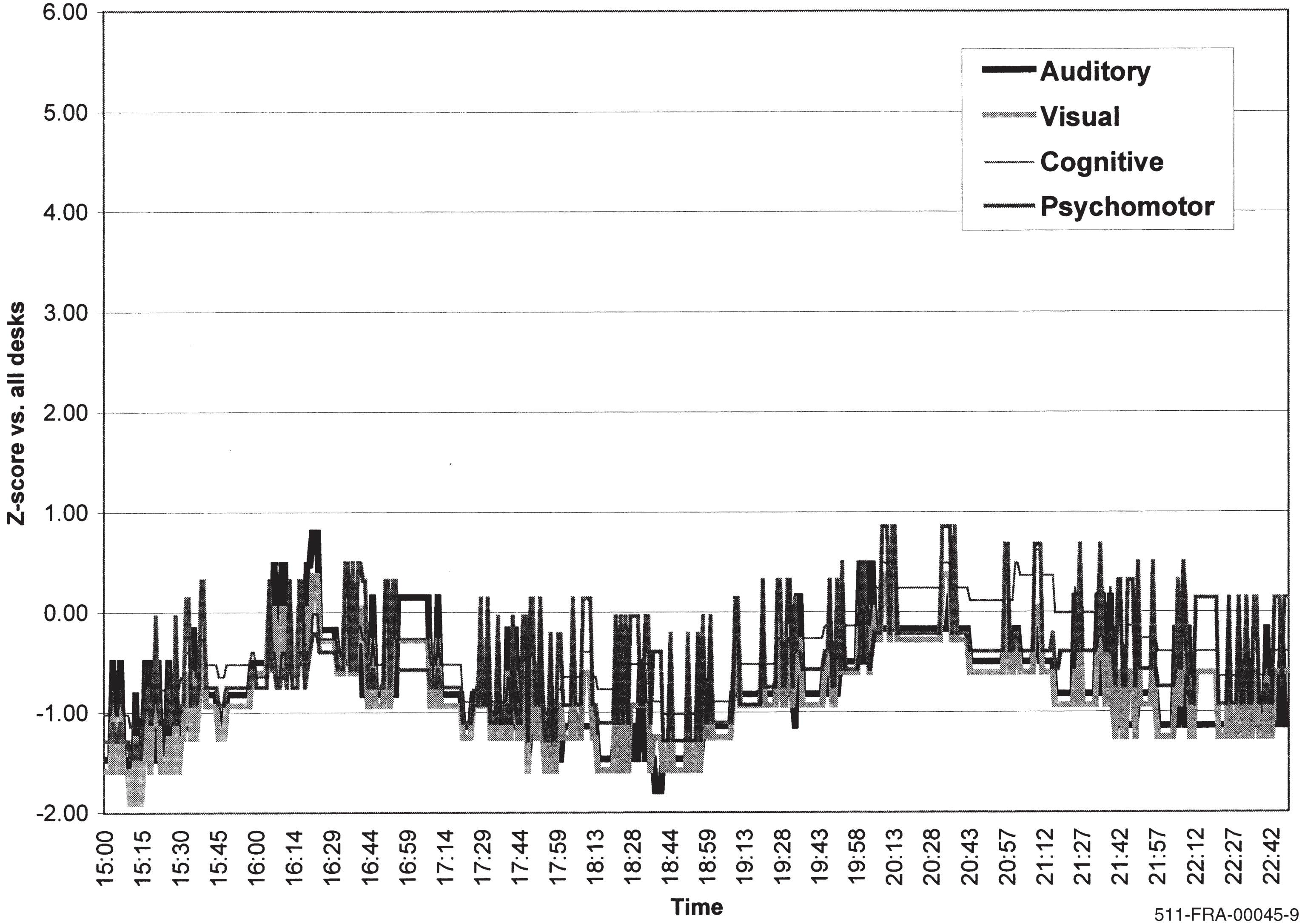


Figure 14. Freight operations, Saturday, evening shift, Desk D

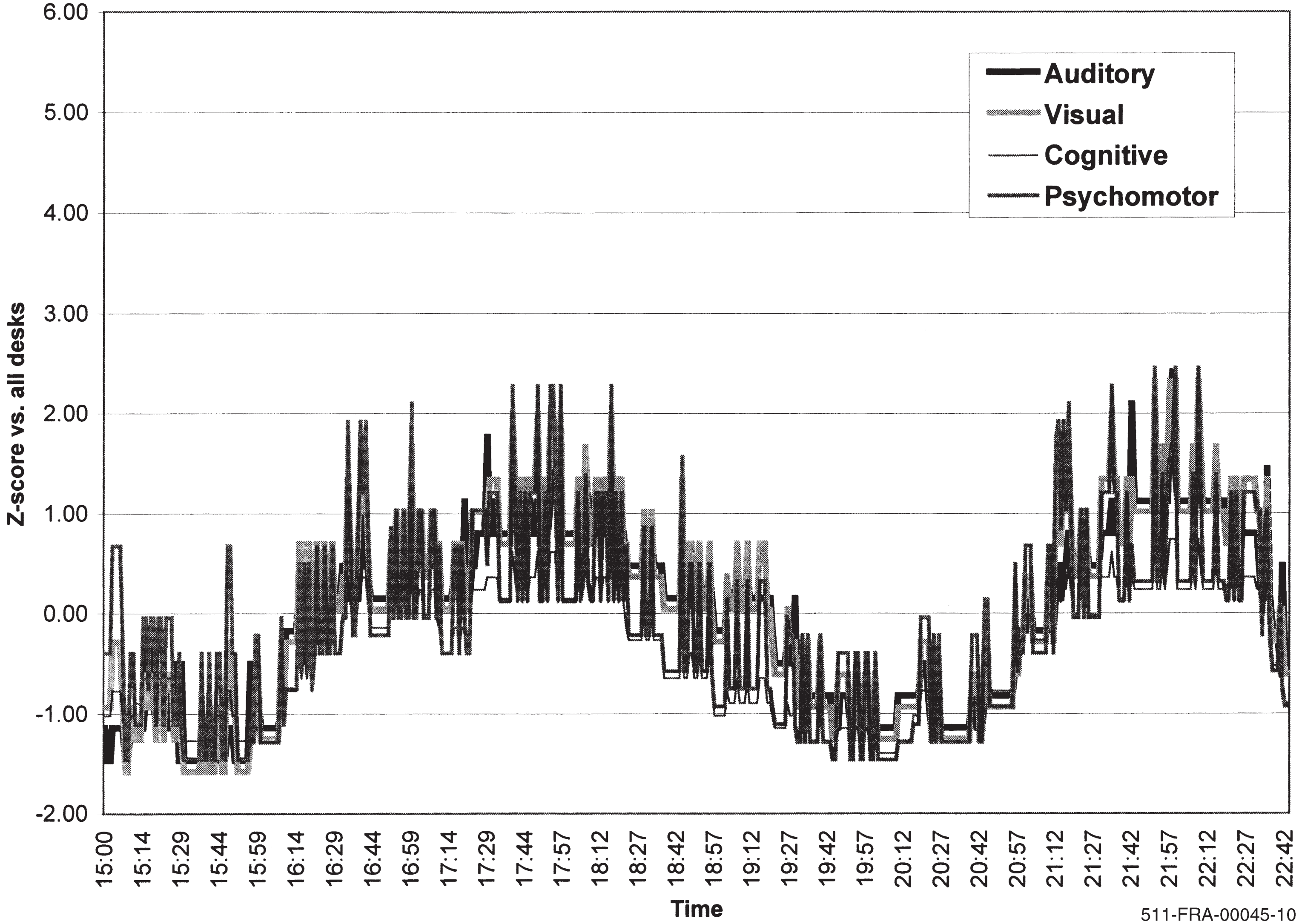


Figure 15.   Freight operations, Tuesday, evening shift, Desk G

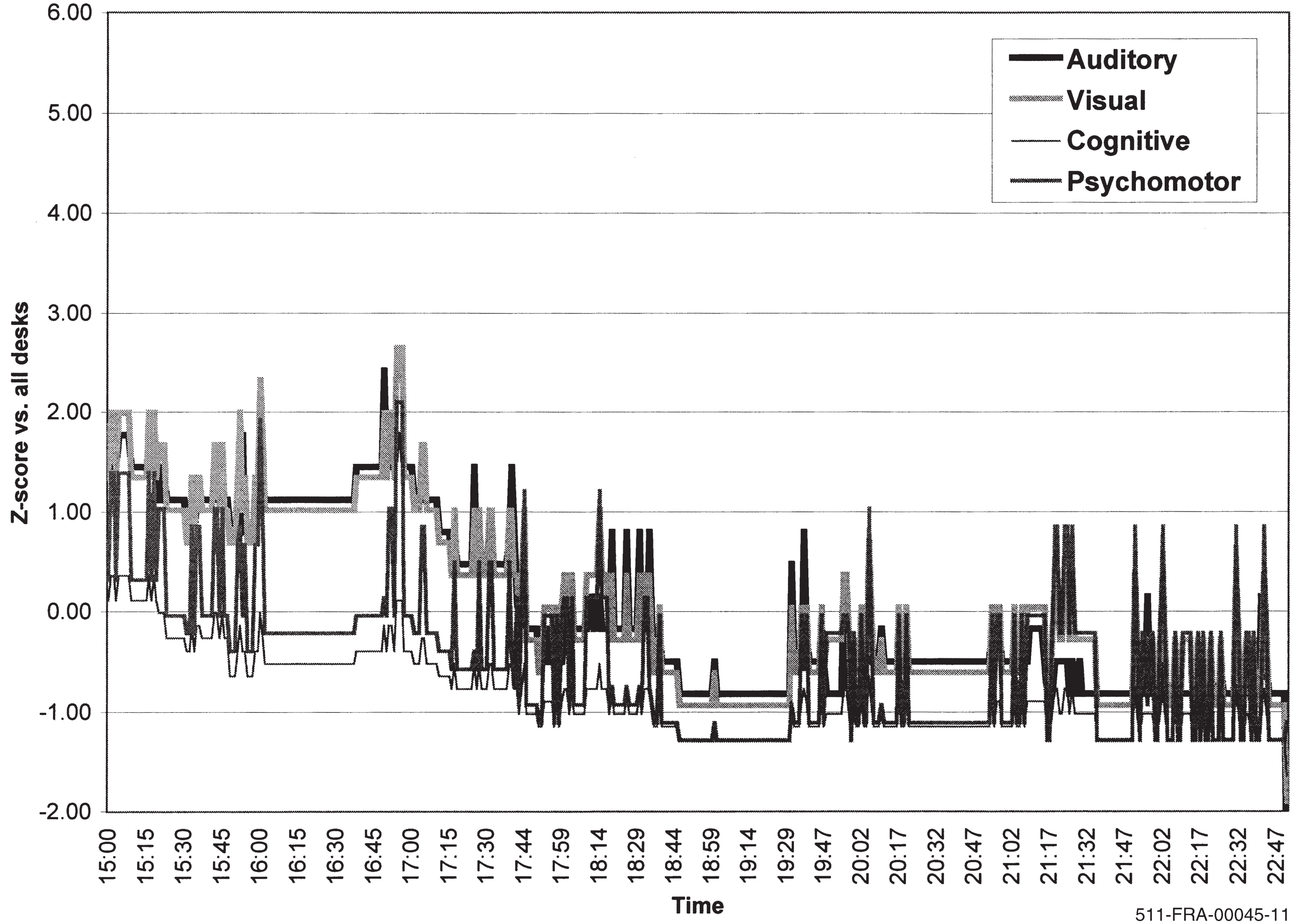


Figure 16. Freight operations, Monday, evening shift, Desk F

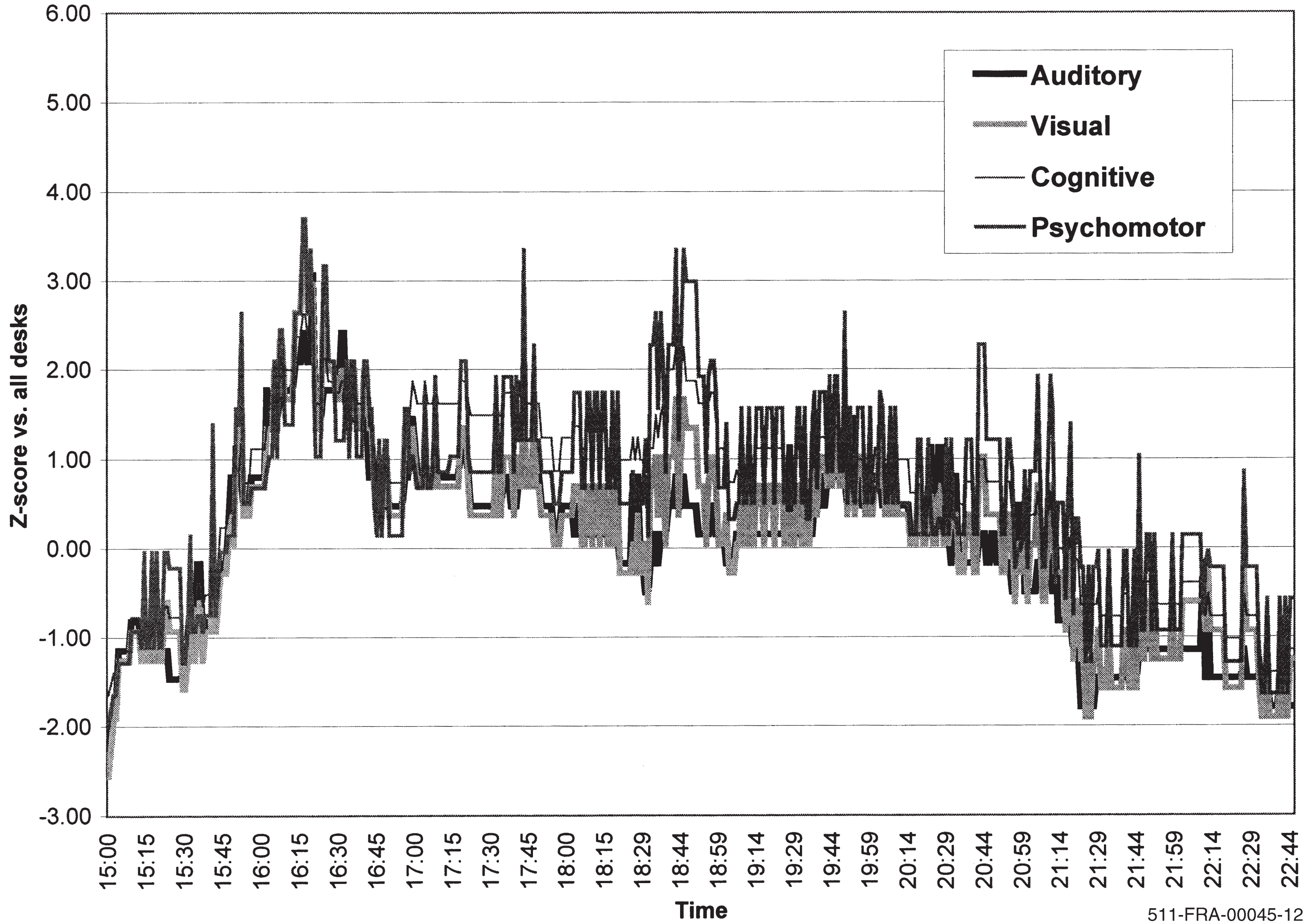


Figure 17. Freight operations, Sunday, evening shift, Desk E

Table 15. mTAWL overload rating by desk

Type of Operation	Day	Work Shift	Desk Observed	% of Time Overloaded within Shift	Minutes of Time Overloaded within Shift	% of Time Overloaded Across Desks	Minutes of Time Overloaded Across Desks
Freight	Tuesday	1	A	3.4	16	3.8	18
	Wednesday	1	B	1.1	5	1.8	9
	Thursday	1	C	0.9	4	0.2	1
	Friday	1	A	1.1	5	17.5	84
	Sunday	2	D	0.0	0	0.0	0
	Monday	2	E	2.1	10	3.2	15
	Tuesday	2	F	4.9	24	0.6	3
	Wednesday	2	G	2.6	12	0.9	4
	Thursday	3	D	8.3	40	0.0	0
	Friday	3	A	3.3	16	0.0	0
Passenger	Monday	1	A	4.5	22	0.0	0
	Tuesday	1	B	0.0	0	0.0	0
	Wednesday	1	C	6.7	32	21.0	101
	Thursday	1	B	4.0	19	0.2	1
	Friday	1	D	1.0	5	1.3	6
	Saturday	1	E	2.5	12	0.0	0
	Sunday	1	E	2.6	12	0.0	0
	Monday	1	F	2.1	10	4.2	20
	Tuesday	1	A	7.5	36	0.0	0
	Wednesday	2	B	0.9	4	0.0	0
	Thursday	2	D	0.0	0	0.0	0
	Friday	3	G	4.5	22	0.0	0

from each chart the total percent of time the dispatcher was in an overload state, defined as having at least two channels that are two or more standard deviations above the mean. These values are computed using two separate approaches. The first approach involved standardizing scores individually by desk. This value was then used to determine the degree of workload variability that occurs at that desk during a shift. The second approach involved computing standardized scores using the entire data set collected at the center. This second approach allowed comparisons to be made across desks at the same center to determine if there was an even distribution of workload, or if particular desks are overly burdened relative to other desks that were observed at the dispatching center.

As can be seen in Table 15, each dispatching center contained one desk that appears to overload the dispatcher for a relatively high percentage of the time. For example, day shift desk A at the freight operation and day shift Desk C at the passenger operation produced an overload state for 84 min and 101 min, respectively. The next most overloaded desk at both sites was in an overload state for less than 5 percent of the shift (i.e., less than 30 min). The overload state was likely due to the shift, desk and day of the week in which the desks were observed. Within-shift variability of a single desk does not necessarily lead to similar results as obtained from workload comparisons across desks. Day shift Desk A at the freight operations was overloaded 17.5 percent of the time but only 1.1 percent of the time when compared within shift to itself. This means that the dispatcher is generally experiencing a high workload level throughout the shift. The chart confirms this. Conversely, night shift Desk D at the same operation did not experience any overload when comparing between desks but does show an overload 8.4 percent of the time for the within-desk (i.e., several dispatchers working the same desk on different days) comparison. This indicated a greater range or higher variability in workload over the shift than what was experienced in Desk A.

While the mTAWL data appear to indicate some overload states, it was not possible to collect dispatcher performance (i.e., benchmark) data to validate these conclusions. In fact, dispatcher performance may be affected by high levels of overload, high levels of underload, or by high variability in workload. Although performance data were not collected, the mTAWL data do appear to provide diagnostic information about the individual desks that were observed. In particular, the mTAWL metric, normalized number of minutes overloaded, provides information about the pattern of workload at a center's dispatcher desks, and whether workload is evenly distributed across desks during a given shift within a dispatching center. In relation to this study, it appears that at least one desk per study site may be experiencing excessive workload.

5.3.2 Self-Report Workload

A 2X3X4 ANOVA was computed comparing separately the numbers of trains and other track users reported being handled by the dispatchers between the study sites, across the three shifts, and the two-hour intervals within each shift. (See Table 16 for a summary of the results.) Independent t-tests were also used to compare field sites by shift. (See Table 17.) A more conservative significance level, $p<0.01$, is used to control for Type II error. Of the significant within-shift differences found between these two field sites, typically it was the passenger operation that reported a greater number of trains and track users. While the majority of these differences occurred during the night shift, several differences were also found during the day and afternoon shifts. These data are plotted in Figure 18 through Figure 20.

As is shown in these figures, the number of trains handled at the freight operation, both across shift and across a 24 hr period, is more evenly distributed than at the passenger operation. This was expected as the passenger operation was subject to rush hour surges in the morning and evening commuting hours. The number of track users also fluctuated more for the passenger operation since the territory that these dispatchers controlled was undergoing a lot of heavy maintenance. The freight operation appears to dispatch consistently more trains than the passenger operation except during daily work commute hours. The passenger operation consistently dispatched more track users throughout a 24 hr period.

Shift effects, tested by repeated measures ANOVA, were calculated separately for each study site. Analyses were separated by field site for two reasons. First, the ANOVAs and *t*-tests that are reported in Tables 16 and 17 already indicated a significant difference between the two operations. Secondly, there was an unusually large amount of night MOW activity with the

Table 16. ANOVA evaluation of effect of type of operation, shift and time into shift on workload measures

Comparison	F	DF	p
Number of Trains			
Site	1.58	1, 1356	0.208
Shift	18.58	2, 1356	0.000
Time into Shift (TIS)	5.994	3, 1356	0.000
Site X Shift	11.58	2, 1356	0.000
Site X TIS	3.00	3, 1356	0.030
Shift X TIS	1.66	6, 1356	0.127
Site X Shift X TIS	1.74	6, 1356	0.109
Number of Track Users			
Site	31.36	1, 1356	0.000
Shift	107.92	2, 1356	0.000
Time into Shift (TIS)	0.640	3, 1356	0.589
Site X Shift	0.714	2, 1356	0.490
Site X TIS	0.191	3, 1356	0.903
Shift X TIS	1.995	6, 1356	0.063
Site X Shift X TIS	1.7	6, 1356	0.117

passenger operation due to the upgrading of the Northeast corridor, a short-term situation not relevant to the freight operation. Results of the repeated measures ANOVA are reported in Table˚18. Significant time-into-shift effects were found for both the number of trains and the number of other track users dispatched in the freight operation, but only the number of trains dispatched in the passenger operation. A simple shift effect was found in all cases, indicating that the number of trains and other track users dispatched varied by shift. Table E-1 in Appendix E further examines these shift effects through a series of Tukey Honestly Significant Difference (HSD) post hoc comparisons. The large number of differences between the shifts suggests that each shift should be considered separately.

5.3.3 Subjective Ratings

ANOVA was used to evaluate the effects of field site, shift and time-into-shift on subjective workload, stress and fatigue. Results of these analyses may be found in Table 19. Table E-2 in Appendix E presents the mean subjective workload, stress, and fatigue ratings by study site for each shift and data collection period. This table also includes baseline subjective ratings that were collected just prior to the onset of each work period. The first shift dispatchers at the passenger operation started off with higher subjective ratings of workload, stress, and fatigue than the freight operation. This may be due to the anticipation of the rush hour dispatching they are about to handle, a difference in the organizational climate of the dispatching centers, the difference between an urban and a rural setting, or the difference in commute times to work or ease of commute. These workload, stress and fatigue differences continue through the morning

Table 17. Comparison of average number of trains and track users handled every 2 hr, separated by shift and field site

Shift	Site	M Trains 2 hr	M Users 2 hr	M Trains 4 hr	M Users 4 hr	M Trains 6 hr	M Users 6 hr	M Trains 8 hr	M Users 8 hr	M Trains Overall	M Users Overall
Day	Freight	7.96	4.10	8.48	5.45	8.32	5.29	8.49	5.07	8.31	4.98
	Pass.	**13.23***	7.16	10.56	9.43	8.71	6.70	8.23	5.55	**10.17***	**7.16***
Evening	Freight	11.02	2.83	10.71	1.07	9.43	0.66	9.79	0.5	10.20	1.24
	Pass.	11.51	2.76	9.89	1.83	8.78	**2.57***	7.58	**3.08***	9.46	**2.56***
Night	Freight	8.70	0.17	**7.93***	0.18	**7.6***	0.28	8.02	0.57	**8.11***	0.30
	Pass.	7.63	**2.72***	2.88	1.81	3.55	**2.35***	7.81	**1.66***	5.51	**2.13***

*Indicates $p<0.01$ level of significance on an independent sample t-test.

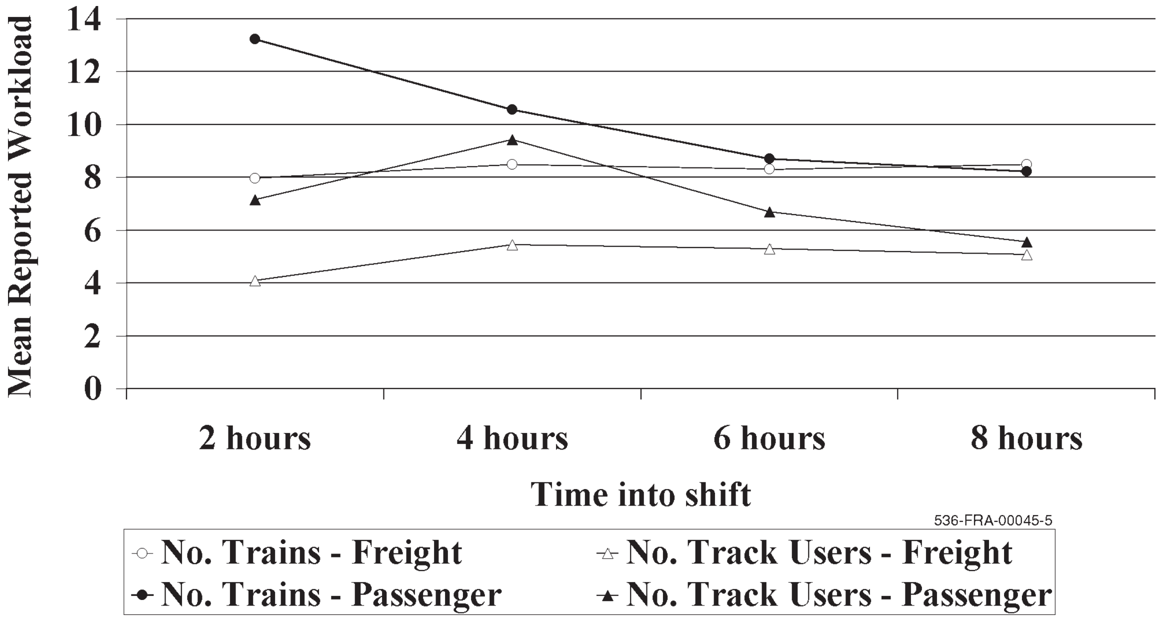


Figure 18. Workload comparisons for day shift by type of operation and time into shift

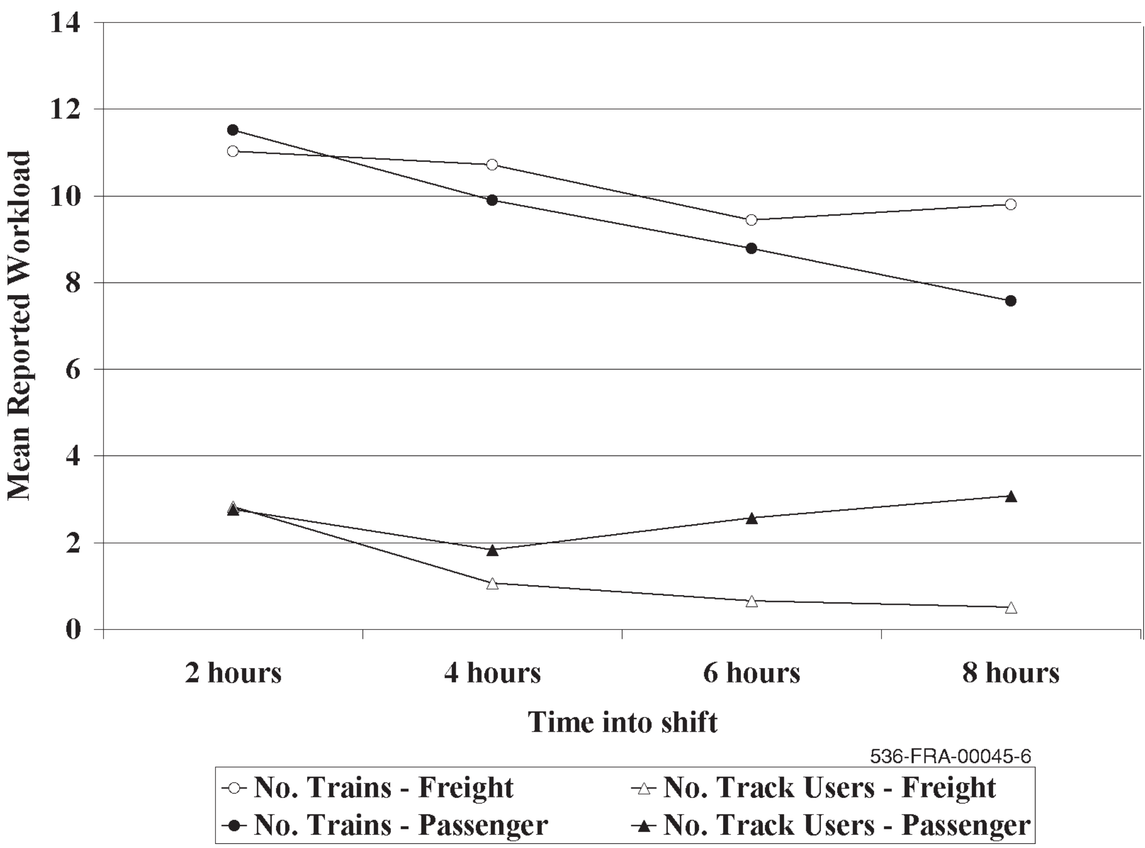


Figure 19. Workload comparisons for evening shift by type of operation and time into shift

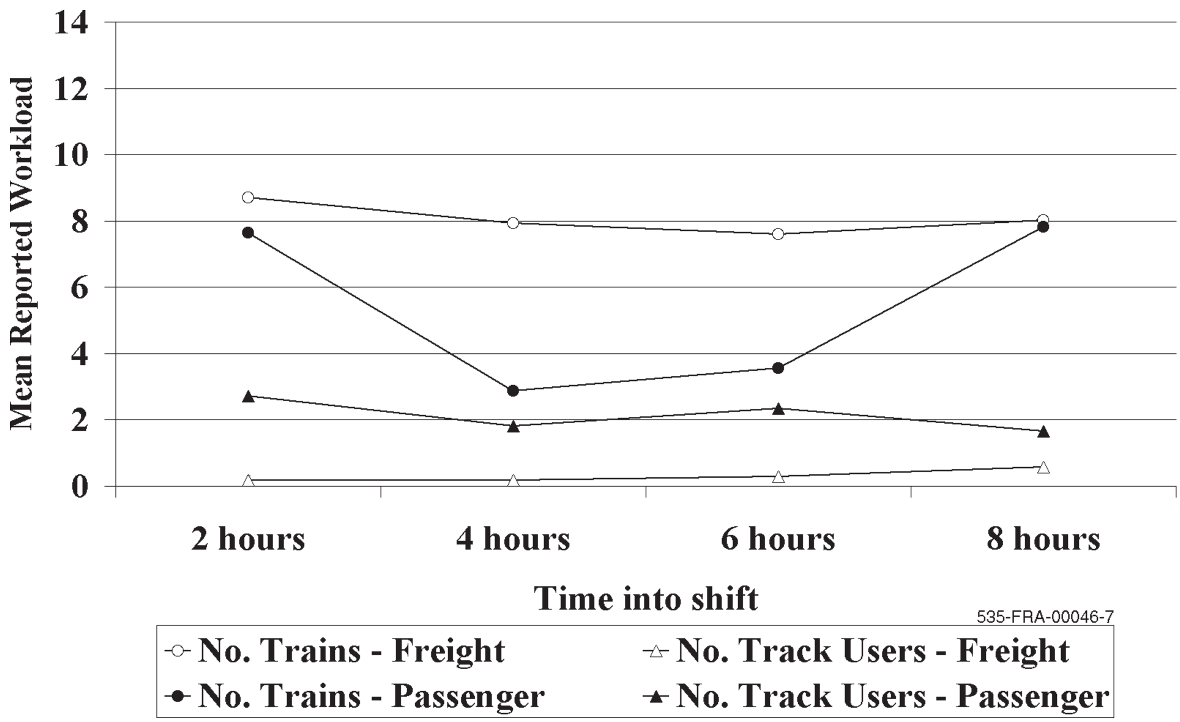


Figure 20. Workload comparisons for night shift by type of operation and time into shift

Table 18. Repeated measures analysis on number of trains and track users over shift

Type of Operation	Workload Measure	Time into Shift Effect	Shift Effect
Freight	Number of Trains	**$F(3,558)=4.225$***	**$F(6,558)=2.818$***
	Number of Track Users	**$F(3,567)=2.794$***	**$F(6,567)=21.537$***
Passenger	Number of Trains	**$F(3,441)=20.599$***	**$F(6,441)=7.592$***
	Number of Track Users	$F(3,405)=1.24$	**$F(6,405)=3.685$***

*Indicates $p<0.05$ level of significance.

rush hour, after which they fall off except feelings of fatigue, which persist through the 6 hr mark into the shift. The only difference found with the evening shift was that the dispatchers at the passenger operation reported higher perceived workload 2 hr into shift. This coincides with the evening rush hour period.

The freight operation dispatchers reported significantly higher levels of subjective workload and stress at points during the night shift than did the passenger operation dispatchers. Again, this fits with the pattern of workload data based on number of trains and other track users. Interestingly, the only fatigue difference for this shift came from the passenger operation dispatchers after 2 hr on shift. Subjective workload, stress, and fatigue data are plotted in

Table 19. ANOVA evaluation of effect of type of operation, shift and time-on-shift on workload, stress and fatigue

Comparison	F	DF	*p*
Subjective Workload			
Type of Operation	4.289	1, 1737	0.038
Shift	4.111	2, 1737	0.017
Time into Shift (TIS)	1.231	4, 1737	0.296
Site X Shift	28.508	2, 1737	0.000
Site X TIS	1.158	4, 1737	0.328
Shift X TIS	5.484	8, 1737	0.000
Site X Shift X TIS	1.860	8, 1737	0.062
Subjective Stress			
Type of Operation	3.224	1, 1737	0.073
Shift	5.923	2, 1737	0.003
Time into Shift (TIS)	9.135	4, 1737	0.000
Site X Shift	9.695	2, 1737	0.000
Site X TIS	0.953	4, 1737	0.432
Shift X TIS	3.730	8, 1737	0.000
Site X Shift X TIS	1.333	8, 1737	0.222
Subjective Fatigue			
Type of Operation	36.912	1, 1737	0.000
Shift	57.183	2, 1737	0.000
Time into Shift (TIS)	35.582	4, 1737	0.000
Site X Shift	4.785	2, 1737	0.008
Site X TIS	0.124	4, 1737	0.974
Shift X TIS	0.556	8, 1737	0.814
Site X Shift X TIS	0.450	8, 1737	0.891

Figure 21 through Figure 23. In general, subjective workload seems to follow traffic volume, that is, the number of trains and other track users. Feelings of stress and fatigue appear to linearly increase throughout the shift, somewhat independent of perceived workload, study site, and shift worked. The one exception to this was that level of perceived stress seemed to closely follow perceived workload for passenger operation dispatchers on the night shift, both being very low. Subjective fatigue always increased over the course of the shift, at least doubling during the day and evening shifts, and nearly doubling during the night shift for dispatchers at both study sites.

Repeated measures ANOVAs calculated separately for each study site on participant self ratings are reported in Table 20. Significant main effects, in terms of hours into shift, were found on perceived workload, stress and fatigue for passenger operation dispatchers, while significant main effects were found only for perceived stress and fatigue for the dispatchers at

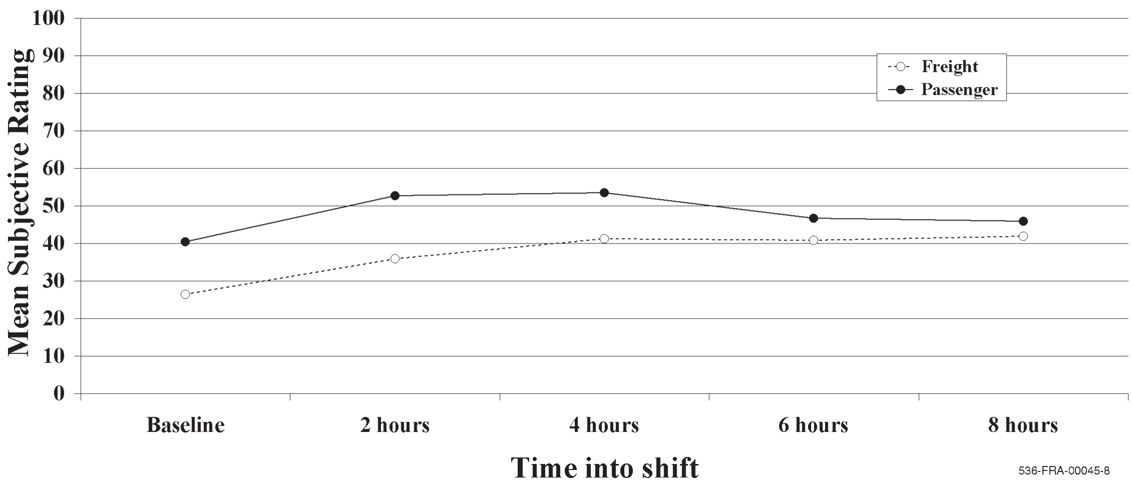


Figure 21a. Subjective ratings comparison of workload for the day shift

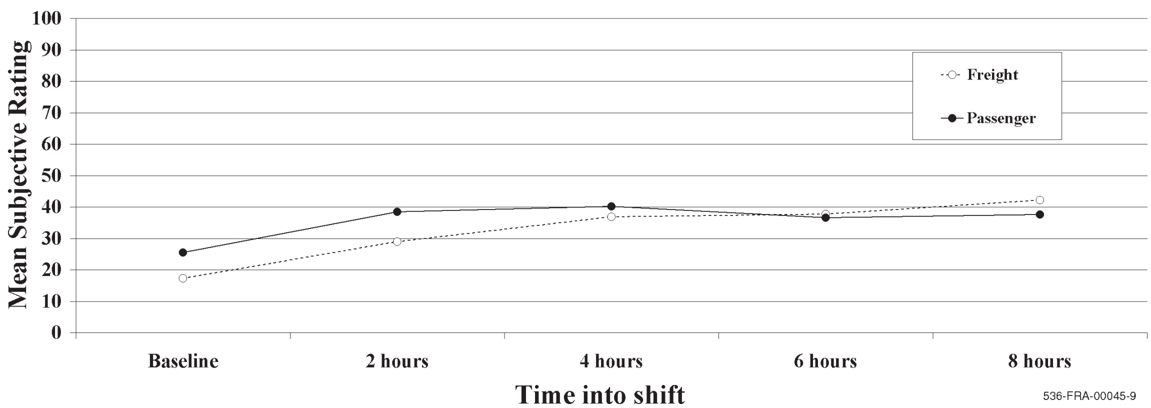


Figure 21b. Subjective ratings comparison of stress for the day shift

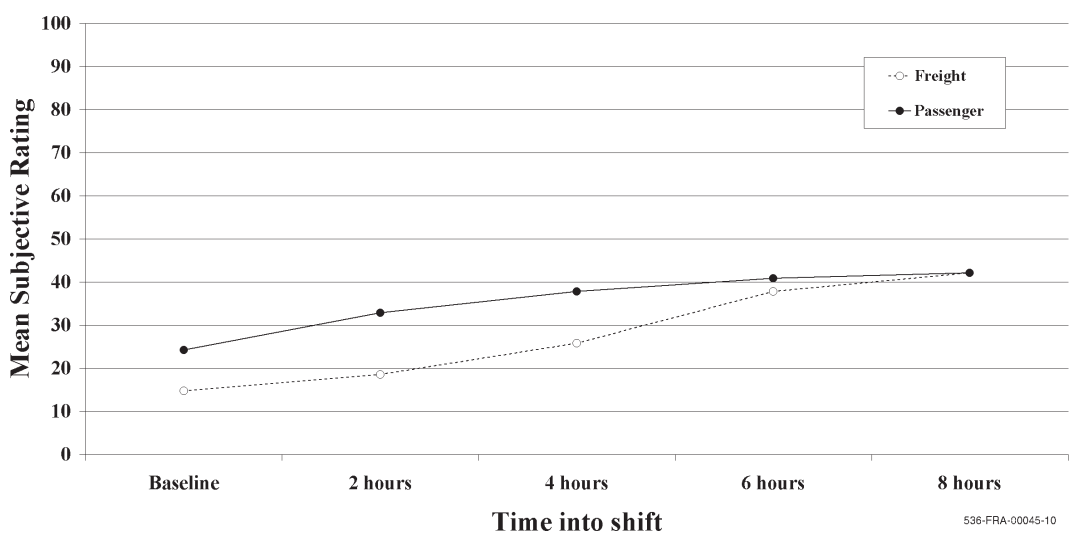


Figure 21c. Subjective ratings comparison of fatigue for the day shift

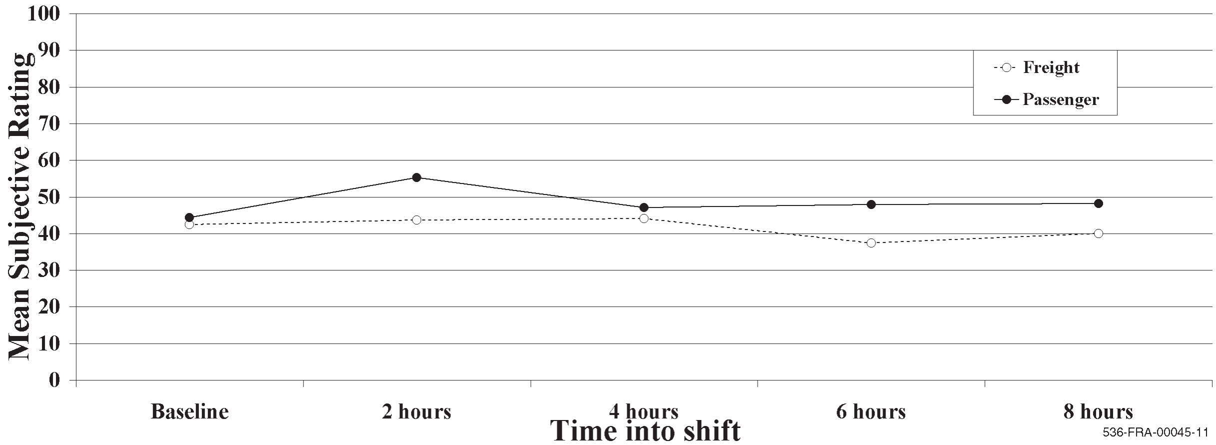


Figure 22a. Subjective ratings comparison of workload for the evening shift

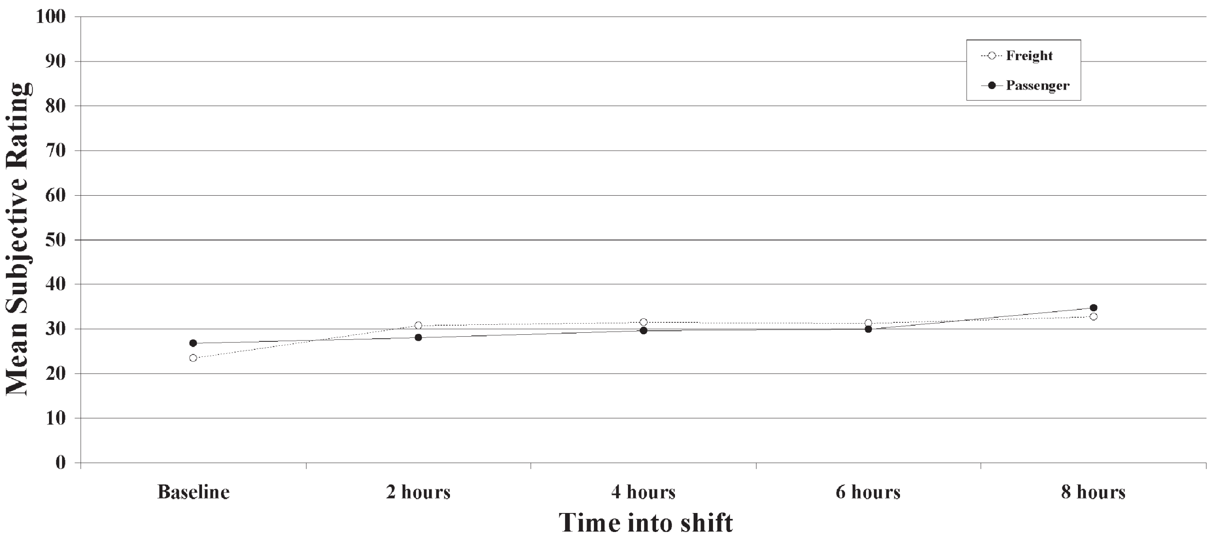


Figure 22b. Subjective ratings comparison of stress for the evening shift

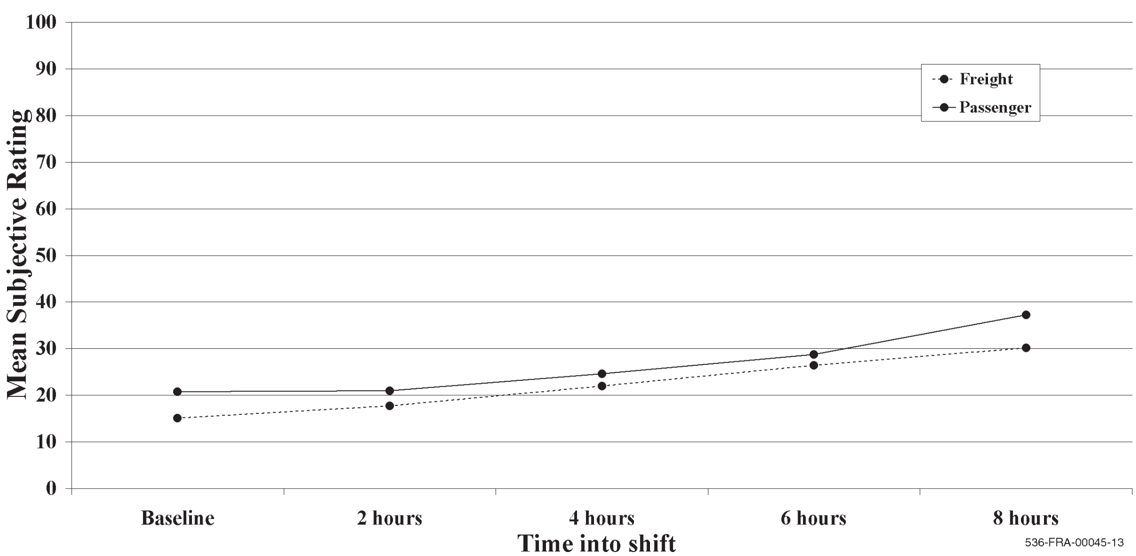


Figure 22c. Subjective ratings comparison of fatigue for the evening shift

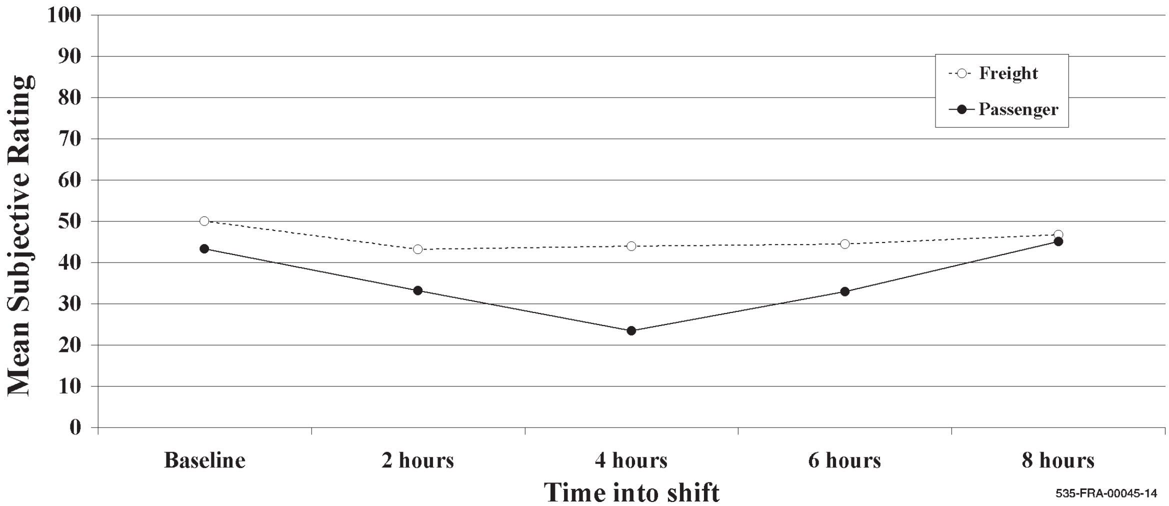


Figure 23a. Subjective ratings comparison of workload by type of operation for the night shift

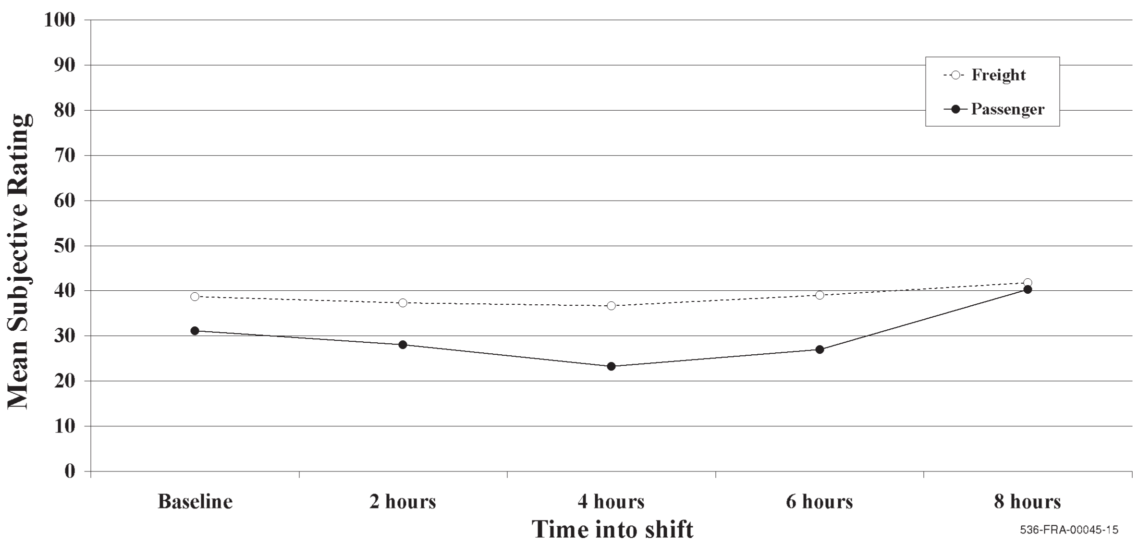


Figure 23b. Subjective ratings comparison of stress by type of operation for the night shift

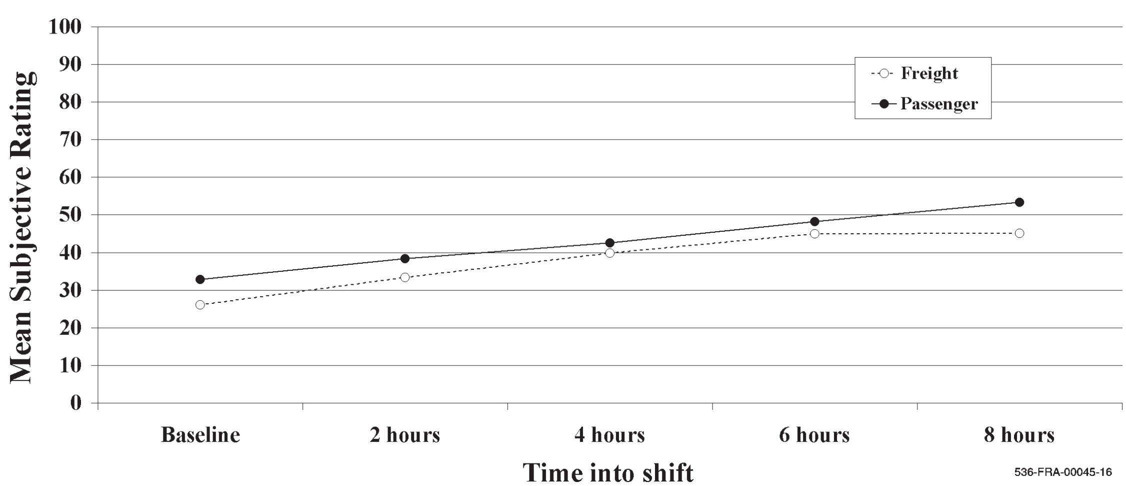


Figure 23c. Subjective ratings comparison of fatigue by type of operation for the night shift

Table 20. Repeated measures analysis on perceived workload, stress and fatigue over shift by type of operation

Type of Operation	Perception	Hours into Shift	Shift Worked
Freight	Perceived Workload	$F(4,732)=1.840$	**$F(8,732)=7.966$***
	Perceived Stress	**$F(4,732)=20.221$***	**$F(8,732)=7.697$***
	Perceived Fatigue	**$F(4,732)=66.835$***	$F(8,732)=0.921$
Passenger	Perceived Workload	**$F(4,560)=2.536$***	**$F(8,560)=6.928$***
	Perceived Stress	**$F(4,556)=5.057$***	**$F(8,556)=5.057$***
	Perceived Fatigue	**$F(4,556)=39.757$***	$F(8,556)=0.177$

*Indicates $p<0.05$ level of significance.

the freight operation. A simple shift effect was found for only workload and stress ratings at both operations; levels of perceived fatigue were not found to be associated with shift worked.

These shift effects were further examined through a series of Tukey HSD post hoc comparisons. Complete post hoc results are in Table E-3 in Appendix E. The pattern of results for passenger operation dispatchers indicates that all three shifts start off with relatively high ratings of subjective workload, stress and fatigue, but that the slope of these functions differs across time, though having a similar asymptote. Night shift dispatcher ratings of perceived fatigue increase more quickly over time than those of the day and evening shift dispatchers. Otherwise, workload and stress seem to increase more quickly during the day shift than the afternoon shift, and the night shift values seem to be more closely associated with the measures of reported workload. Results for the freight operation dispatchers were similar in that fatigue accumulated more quickly during the night shift than the day and evening shifts. There was not as clear a difference with workload and stress for the freight operation dispatchers, however. This is likely a result of the consistent workload that they experienced throughout the course of the day.

Correlations were computed between both number of reported trains and track users and perceived workload, stress and fatigue. These computations were done for each data collection point within each shift separately for each field location. (Summary statistics are in Table 21 and the complete correlation matrix is located in Appendix E.) Perceived workload was moderately associated with reported number of trains dispatched, regardless of shift or study location. These correlations were found to be significant, and therefore reliable, for the majority (i.e., 92˚percent) of the computations. Perceived stress was also related to the number of trains dispatched, especially at the freight operation where 92˚percent of the coefficients were significant and reliable, compared to 42˚percent at the passenger operation. Subjective fatigue had the lowest association with the number of trains and other track users dispatched, particularly among the dispatchers at the passenger operation. In fact, the strength of the reliable associations between number of track users and subjective workload, stress and fatigue ratings was found to be half of what it was when using number of trains handled by the passenger operations dispatcher. This

Table 21. Percent significant correlations

Type of Operation	Shift	Number of Trains by Workload	Number of Trains by Stress	Number of Trains by Fatigue	Number of Track Users by Workload	Number of Track Users by Stress	Number of Track Users by Fatigue
Overall		92	67	29	50	54	25
Freight	Day	100	75	25	100	100	75
	Evening	100	100	75	50	100	25
	Night	100	100	50	50	25	0
Passenger	Day	75	50	0	0	0	0
	Evening	75	25	25	25	50	50
	Night	100	50	0	75	50	0

difference in strength of association was less pronounced for the freight operations dispatcher. These results suggest that feelings of fatigue are not merely a function of workload as defined by number of trains and other track users handled. Furthermore, the subjective ratings of the dispatchers at the passenger operation appear to be affected differently by trains than by other track users; this effect is less pronounced in the freight dispatchers subjective ratings.

Figure 24 and Figure 25 present Form D count data along with subjective workload, stress and fatigue data over 24 hr for the two study sites. The graph for the freight operation shows a decline in the number of Form Ds over the course of the three shifts while the subjective measures either increased or remained constant. Few individual correlation coefficients computed between the bi-hourly number of Form Ds and subjective ratings recorded at either operation reached statistical significance. Those coefficients reaching statistical significance tended to be at the 0.4 level or lower, with the exception of the passenger operation dispatchers whose subjective workload reports correlated with Form Ds at levels approaching 0.6. Multiple regression analyses were also performed on these data. Separate analyses were computed for each field site and for each shift within a field site. The results, presented in Table 22, suggest that while there is a reliable linear relationship between number of Form Ds and route blocks handled and subjective ratings, their actual impact on the subjective ratings is minute. Route block count data from the freight operation were tested in a similar manner (these data were not available from the passenger site due to operational differences). The pattern of results from these data were similar to that found with the Form D data (see Figure 26 and Table 23).

5.3.4 Summary

It appears that the available count data (i.e., number of trains, track users, Form Ds and route blocks) were not particularly good indictors of subjective workload, stress and fatigue when used individually. These data were also combined through stepwise regression analyses to determine if some combination of these variables would lead to better prediction. Table 24 presents the results of this analysis, which was broken down by field site and subjective variable. While all full models included number of trains handled, two of the three freight models also included total number of track users; all three passenger operations models included number of Form Ds

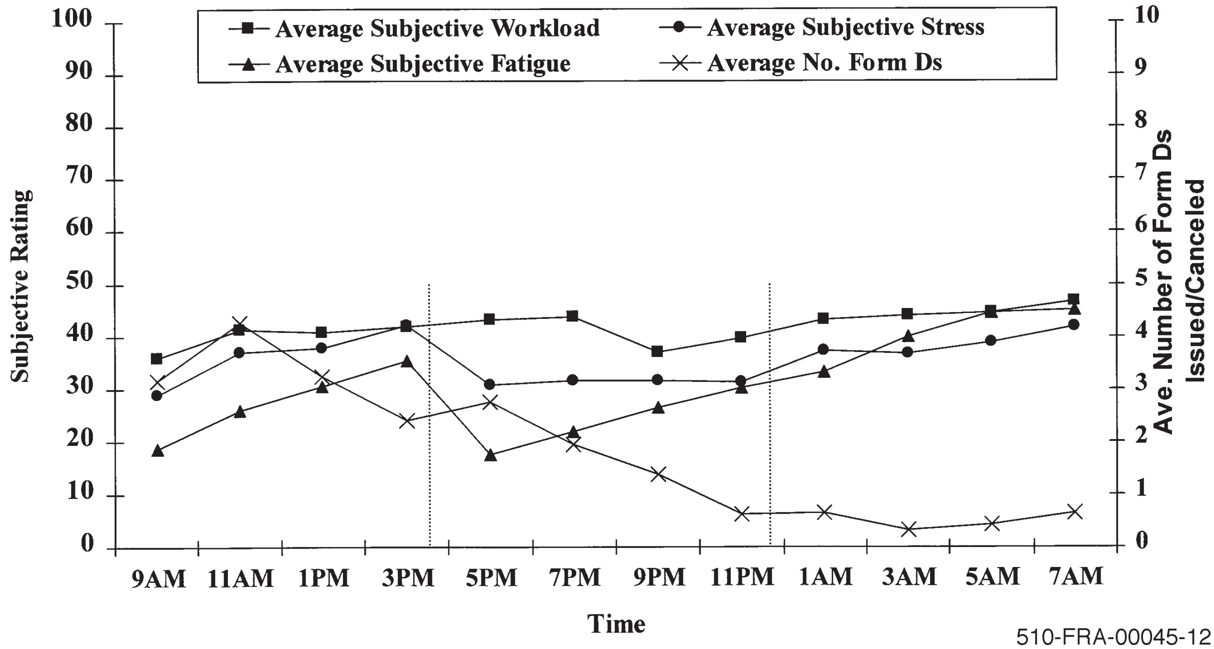


Figure 24. Freight Form Ds and subjective ratings across 24 hr

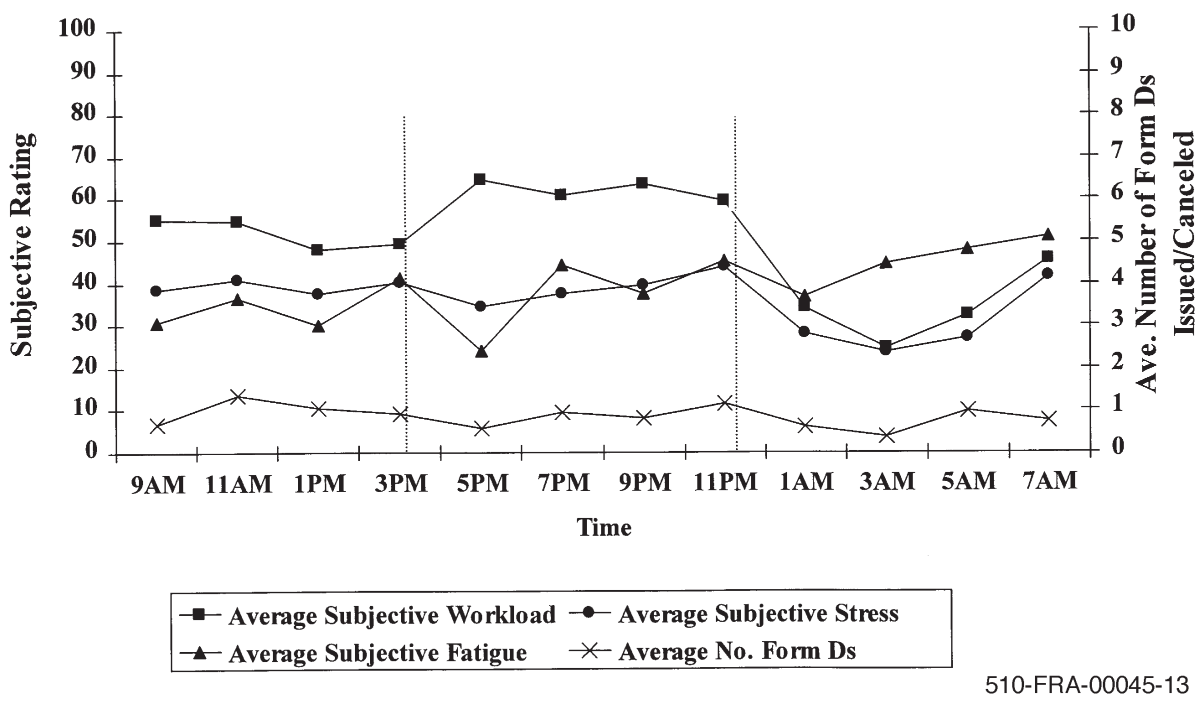


Figure 25. Passenger Form Ds and subjective ratings across 24 hr

Table 22. Multiple regression analysis results by field location and shift for predicting subjective ratings as a function of number of Form Ds completed

Field Location	Shift	Predicted Variable	R^2	*df*	*F*	*p*
Freight	Day	Subjective Workload	0.059	1,262	17.553	0.000
		Subjective Stress	0.087	1,262	26.032	0.000
		Subjective Fatigue	0.086	1,262	25.754	0.000
	Afternoon	Subjective Workload	0.031	1,268	9.690	0.002
		Subjective Stress	0.038	1,268	11.762	0.001
		Subjective Fatigue	0.017	1,268	4.637	0.032
	Night	Subjective Workload	0.001	1,237	0.816	0.367
		Subjective Stress	0.002	1,237	0.417	0.519
		Subjective Fatigue	0.018	1,237	5.424	0.021
	All	Subjective Workload	0.020	1,771	16.601	0.000
		Subjective Stress	0.038	1,771	31.140	0.000
		Subjective Fatigue	0.005	1,771	4.861	0.028
Passenger	Day	Subjective Workload	0.077	1,333	28.705	0.000
		Subjective Stress	0.061	1,333	22.624	0.000
		Subjective Fatigue	0.004	1,333	2.389	0.123
	Afternoon	Subjective Workload	0.052	1,145	9.032	0.003
		Subjective Stress	0.089	1,144	15.092	0.000
		Subjective Fatigue	0.065	1,145	11.132	0.001
	Night	Subjective Workload	0.122	1,135	19.969	0.000
		Subjective Stress	0.117	1,135	18.962	0.000
		Subjective Fatigue	0.061	1,135	9.892	0.002
	All	Subjective Workload	0.073	1,617	49.349	0.000
		Subjective Stress	0.077	1,616	52.492	0.000
		Subjective Fatigue	0.020	1,617	13.573	0.000

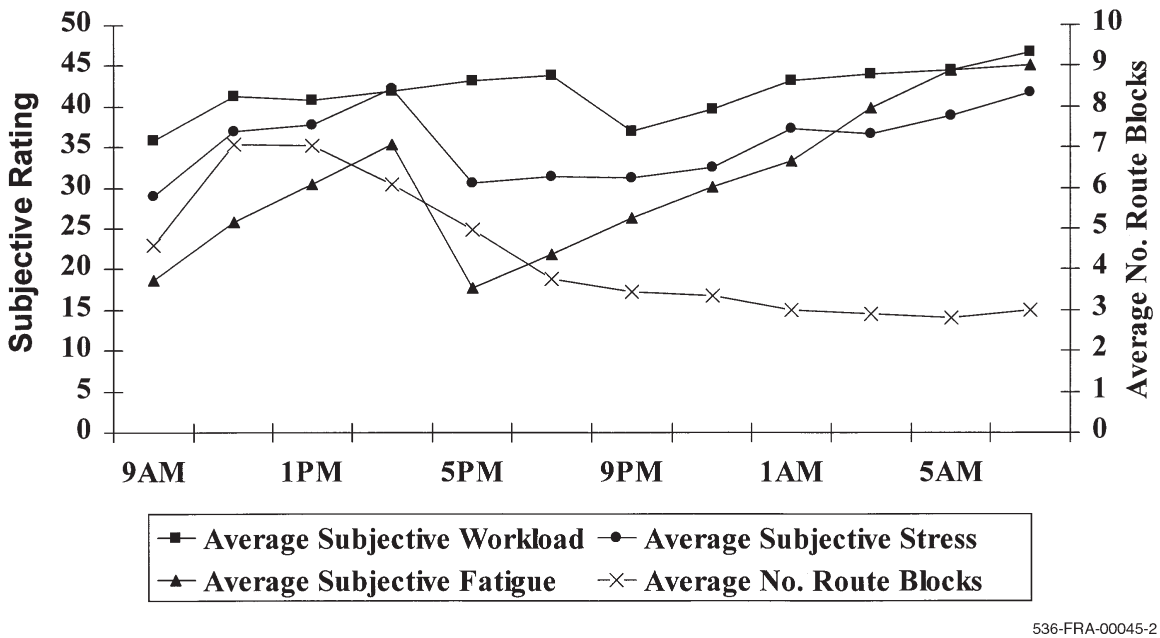


Figure 26. Freight route blocks and subjective ratings across 24 hr

completed. Combining these count variables did lead to significantly better prediction equations, especially for subjective workload, with nearly one-third of the variance explained. Subjective fatigue ratings seem relatively insulated from these count variables.

There are several explanations as to why these count variables did not explain a greater proportion of the variance in these regression equations, especially for subjective workload. First is the frequency of data collection. Perhaps data should have been collected more frequently than every 2 hr (i.e., 2 hr may have been too long of an interval for someone to judge their workload). It is also possible that none of these metrics really tapped the cognitive components of dispatching. Focusing on the visible portion of a dispatcher's job (i.e., activity counts) is perhaps analogous to viewing the tip of an iceberg. That is, 95 percent of what dispatchers do (e.g., planning and decision making) occurs in their heads, and therefore activity counts may not be effective measures since they tap only a small fraction of dispatchers' job activities. Nonetheless, the FAA has developed validated approaches for calculating mental workload of its air traffic controllers based upon the number of aircraft handled. Each type of aircraft and its usage of air space has an associated mental workload value. Therefore, possible workload measures include counting every action taken by the dispatcher or assigning a specific cognitive load to each type of situation the dispatcher might encounter, similar to the mTAWL approach. Further development of this type of approach for railroad dispatcher workload analysis is needed.

The current metrics did provide some useful information, however. Specifically, the mTAWL appears to have potential as a diagnostic metric for evaluating and comparing the workload distribution of the desks at a dispatcher center. This methodology identified one desk at each center where the dispatcher appeared to be in an overload state for a disproportionately

Table 23. Multiple regression analysis results for freight operation by shift for predicting subjective ratings as a function of number of route blocks completed

Shift	Predicted Variable	R^2	*df*	*F*	*p*
Day	Subjective Workload	0.131	1,258	40.106	0.000
	Subjective Stress	0.112	1,258	32.409	0.000
	Subjective Fatigue	0.021	1,258	60.645	0.011
Afternoon	Subjective Workload	0.082	1,268	25.064	0.000
	Subjective Stress	0.068	1,268	20.617	0.000
	Subjective Fatigue	0.002	1,268	1.566	0.212
Night	Subjective Workload	0.000	1,233	0.138	0.711
	Subjective Stress	0.000	1,233	0.261	0.610
	Subjective Fatigue	0.032	1,233	8.706	0.003
All	Subjective Workload	0.043	1,763	35.106	0.000
	Subjective Stress	0.054	1,763	44.671	0.000
	Subjective Fatigue	0.000	1,763	0.043	0.835

Table 24. Full model results of stepwise regression analyses by type of operation and subjective rating using the count variables as predictors

Type of Operation	Subjective Rating	R^2	Predictor Variables*	*df*	*F*	*p*
Freight	Workload	0.274	1,3	2,754	143.873	0.000
	Stress	0.205	3,1,5	3,753	65.805	0.000
	Fatigue	0.015	1	1,755	12.587	0.000
Passenger	Workload	0.292	3,4,1	3,580	81.295	0.000
	Stress	0.133	4,1	2,580	45.526	0.000
	Fatigue	0.020	4,1	2,581	7.084	0.001

*Note: Variables presented in order of loading; 1=number of trains, 2=number of track users, 3=total number of users, 4=number of Form Ds, 5=number of route blocks.

high percentage of the shift. It also appears that subjective ratings of workload are independent of the number of Form Ds, route blocks and number of train and other track users handled. Even after combining all frequency-count workload data together, they account for only 36 percent to 40 percent of the subjective workload variance. This may be due to the "iceberg effect" mentioned above. However, it is important to note that the dispatchers' subjective workload ratings did not indicate excessively high workload. In fact there were only a few instances where mean subjective workload ratings were beyond 50 percent, 100 percent being extreme workload. None of the self-reports went beyond 70 percent. It is thus possible that these dispatchers simply were not experiencing high workload during the study period. However, it is also possible that the dispatchers did not fully understand the directions for completing the subjective workload ratings. A third possible explanation for the low subjective workload ratings could be methodological. Specifically, asking dispatchers to report subjective workload every 2 hr likely forced the dispatchers to average out their workload over this period with the peaks and troughs in workload being smoothed out in the reported ratings. Perhaps a more sensitive measure, or more frequent data collection, is needed to capture the short bursts of workload that likely transpire.

5.4 Stress

Data from three separate instruments were used to evaluate railroad dispatcher stress. The first source was background and exit survey responses. These surveys examined the prevalence and nature of perceived job-related stress. The second source was self-reported subjective ratings of stress that dispatchers completed every 2 hr while on duty. The stress-related self-ratings consisted of the NASA TLX "time pressure" subscale and an overall "stress level" scale. Subjective stress ratings were provided using 100 mm visual analogue scales. The third source of data came from salivary cortisol collected from dispatchers every 2 hr while on duty.

Given the safety-critical nature of the job, the spurts of high workload, especially with passenger operations, and anecdotal information, it was expected that railroad dispatchers were under a great deal of stress, at least periodically. While workload is likely a large contributor to their stress, other factors may influence acute and chronic stress levels. Some possible factors include relationships with co-workers, amount of perceived control over work, and issues pertaining to equipment reliability. This section attempts to both identify the sources of stress perceived by the dispatchers and to determine their level of stress while on duty.

5.4.1 Self-Report Stressors

The background survey asked the dispatchers to report their general level of stress on a series of 24 items using a four-point response scale ranging from "not at all" to "very much." These 24 items covered workload, personality, operating procedures and organizational policy issues. If the dispatcher responded "somewhat" or "very much" to an item, then that item was considered a significant stressor. Table 25 presents all items that were considered to be significant stressors by at least 30 percent of the participants. These items were predominantly workload-related (e.g., amount of work, pace of work, juggling train and engine (T&E) crew and maintenance of way (MOW) demands).

Table 25. Source of stress by type of operation

Factor	Overall	Freight	Passenger
Management Policies	✓		✓✓
Demands of T&E and MOW Crews		✓	
Personality Conflicts with T&E and MOW Crews			✓✓
Amount of Work	✓	✓	
Difficulty of Work		✓	
Pace of Work	✓✓	✓✓✓	
Lack of Control		✓	
Surges in Workload		✓	
Juggling T&E and MOW Needs	✓✓	✓✓✓	
Quality of Workstation and Equipment			✓✓
Communication Problems	✓		✓✓✓
Unnecessary Phone Calls			✓✓
Duplicate Reporting Procedures		✓	
Training New Dispatchers		✓✓	

✓ = 30%
✓✓ = 40%
✓✓✓ = 50+%

By splitting these data up by site, however, two distinct patterns of stressor items emerge. Dispatchers at the freight operation appeared to be primarily concerned with their workload. Specific workload-related stressors that concerned them included difficulty of work, surges in workload, and lack of control. The passenger operation dispatchers, by contrast, found personal instructions and the physical work environment to be their primary stressors. These items included quality of workstation equipment, communication problems and personality conflicts with T&E and MOW crews.

There were very few items on the survey that 50 percent or more of the participants viewed as contributing "somewhat" or "very much" to their stress level. There are several possible explanations for this general lack of agreement over which items contributed to stress. First, it may be simply that there is a wide range of stressors that is considered to be important by the different dispatchers (i.e., large individual differences). If true, this would make it difficult to develop a focused intervention program to alleviate the problem(s), and would point more to individual counseling sessions. Second, it may be that the sum total of stressors considered together has a greater impact on perceived stress for these dispatchers than the magnitude of any one individual stressor. There may also be an interaction between the number of stressors and the magnitude of their perceived import. The sample size was too small to evaluate this alternative, however. Methodological problems also could have contributed to dispatchers' general lack of agreement in identifying sources of stress. First, primary work-related stressors were not included in this list. Second, the four-point response scale may have been too restrictive. And third, the snapshot that was taken using the one-time survey instrument may have been insufficient to capture chronic work-related stressors. That is, dispatchers could have been influenced by any number of unknown factors that affected their one-time responses to the survey. Lastly, it is possible that there was not much workplace stress at the time of the study.

Given the anecdotal data collected prior to and during the data collection phase of the experiment, it seems unlikely that the dispatch office could be considered a low stress environment. For this reason, these same data were regrouped by job tenure. Three groups were utilized, those with less than two years of dispatching experience (n=8), those with between 2 and 5 years of experience (n=6) and those with more than 5 years of experience (n=23). The results of this grouping procedure are shown in Table 26. It becomes immediately apparent that those dispatchers with the lowest job tenure have, as a group, the fewest number of reported common stressors. The middle group, however, has a wide range of common stressors, from management policies, to personality conflicts, to workload. Stressors associated with workload and loss of sleep have the greatest commonality for these dispatchers. While this result may be due to such a small sample size, the low tenure group has a similar size but a different and smaller cluster of common stressors. Those with the most job tenure also seemed to have a wide range of common stressors though only one, "juggling both T&E and MOW needs," reached 50 percent agreement.

Table 26. Sources of stress by job tenure

Factor	Overall	< 2 Years	2 to 5 Years	> 5 Years
Management Policies	✓		✓	✓✓
Demands of T&E and MOW Crews			✓	✓
Personality Conflicts with Crews		✓✓✓	✓	
Personality Conflicts with Other Dispatchers			✓	
Amount of Work	✓		✓✓✓	✓
Pace of Work	✓✓		✓✓✓	✓
Lack of Control				✓
Emergencies			✓✓✓	
Surges in Workload			✓✓✓	✓
Juggling T&E and MOW Needs	✓✓			✓✓✓
Loss of Sleep			✓✓✓	
Physical Environment		✓	✓	
Quality of Workstation and Equipment		✓✓✓		✓
Communication Problems	✓	✓		✓
Unnecessary Phone Calls			✓✓✓	✓
Duplicate Reporting Procedures			✓✓✓	✓

✓ = 30%
✓✓ = 40%
✓✓✓ = 50+%

These data suggest an interaction between dispatching experience and perceived stressors. In fact it looks as if it were an inverted "U" whereby those with less experience have less commonly acknowledged stressors, perhaps due to reduced expectations or job demands. Furthermore, the most experienced dispatchers are in a position to know more of these potential stressors, perhaps due to exposure to them, but are either better able to cope with them, or have a better perspective as to consequences and thus are better able to filter out stressors of less concern. Those in the middle group may be given more responsibility and expectations while at the same time not having the years of experience to know how to cope with the different situations they are expected to confront on a daily basis. More data needs to be collected to determine what relationships exist between particular stressors and type of operation or job tenure groupings.

The passenger operation dispatchers were also given the opportunity at the end of the study to answer the open-ended question, "List your top five contributors to work stress, most stressful first." The nine recurring items listed by the dispatchers are: No rest breaks, number of concurrent tasks, equipment design and failure, emergencies and unexpected events, office environment, personality conflicts, training, management meddling and always working with a different crew.

Most of these items are similar to those items they checked off on the initial survey. These results indicate large individual differences among the dispatchers in what they consider to be stressors, and that these stressors are most pronounced for those with 2 to 5 years of dispatching experience. Only one of the items reported by participants was directly workload-related, number of concurrent tasks handled. One of the most frequently cited stressors was the lack of official rest breaks. Thus stress at these field sites appears to be multivariate in nature, and according to self-report data, it is not completely centered around the work itself.

Figures 27 and 28 present averaged subjective stress data reported by the dispatchers every 2 hr. Subjective feelings of stress rose above a rating of 40 only once during a 24 hr period. In addition, average ratings of subjective time pressure, a contributor to stress, never rose above 55.

While, on average, there does not appear to be a high level of subjective stress among this group of dispatchers, there does appear to be variation between the two types of operations. These subjective stress ratings are intercorrelated with subjective workload ratings. For Figure 27, the 7 am and 9 am data points are significantly different, with the passenger operation dispatchers reporting higher stress levels. This is logical from a workload perspective as the passenger operations dispatchers are engaged in rush-hour traffic during this period of time. The 3 am and 5 am points are also significantly different. This time, it is the freight operation dispatchers who report higher levels of stress, but again they have a higher workload during this period of time. In addition, there is a 0.5 to 0.8 correlation between subjective stress and workload, depending upon the specific data collection time into shift.

Reported stress levels appear to increase through the shift, even as the workload drops off, as it does towards the end of the day shift for the passenger operations dispatchers. This may denote a cumulative stress effect, or a threshold effect whereby once a particular level of stress

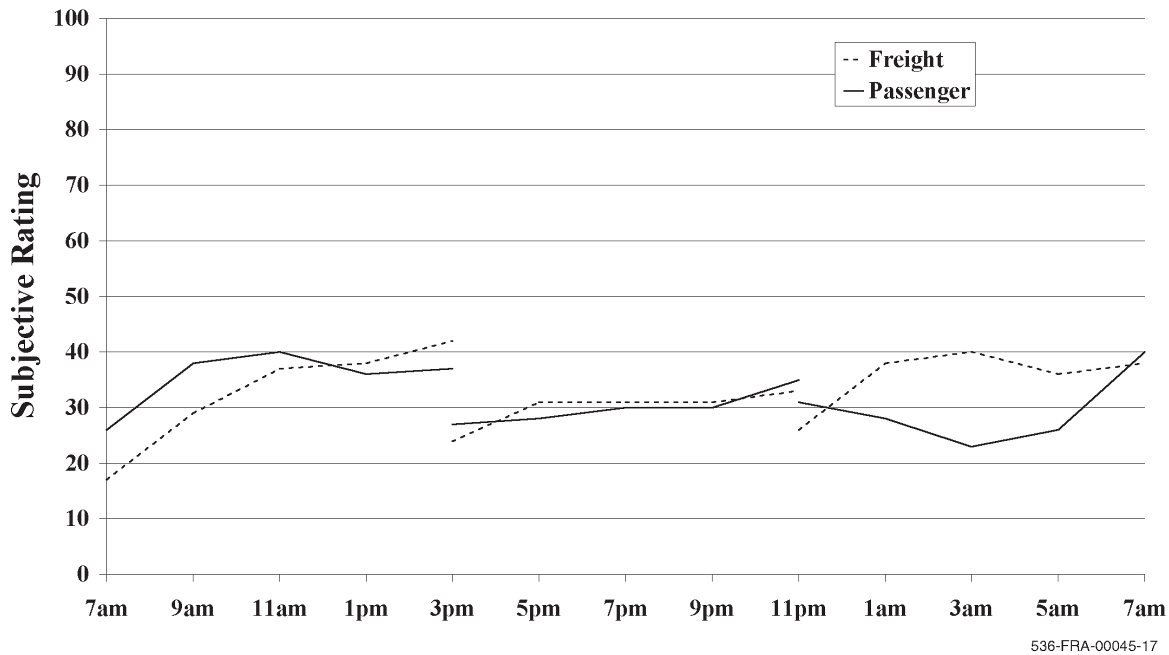


Figure 27. Subjective stress ratings as a function of time of day for passenger and freight operations

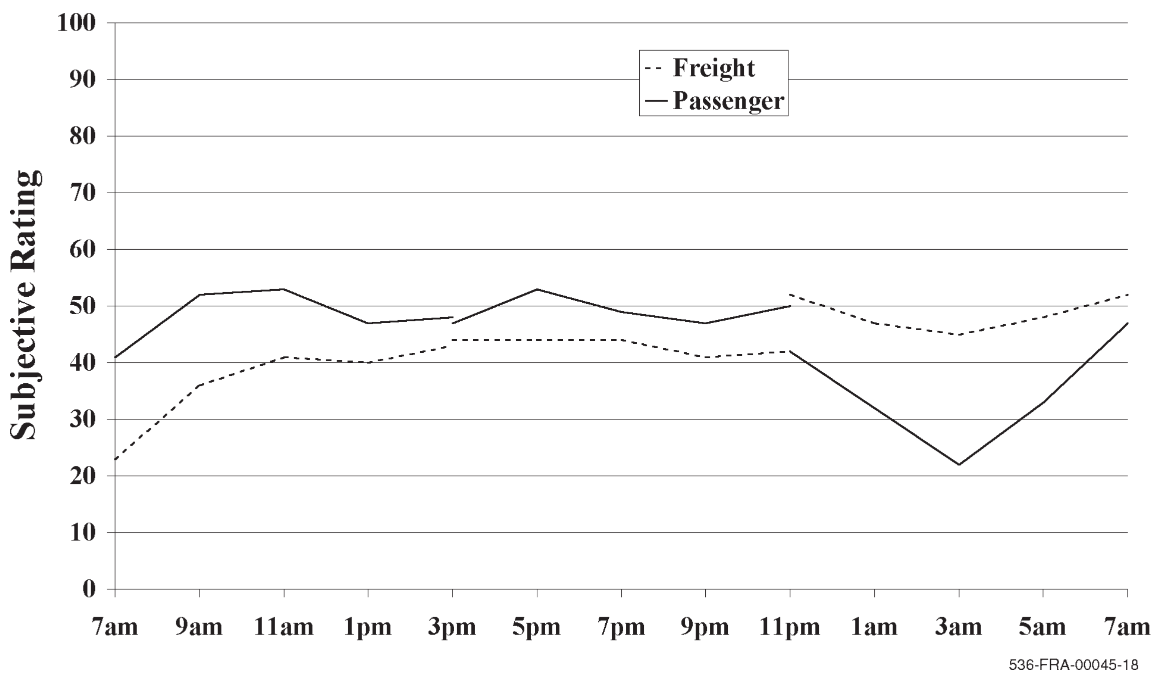


Figure 28. Subjective time pressure ratings as a function of time of day for passenger and freight operations

or arousal has been reached, it is difficult to return to baseline. This idea is evidenced by data from the passenger operation dispatchers working the night shift. Here they started off with a declining workload, which translated into their declining stress ratings. This also indicates that working during the night was not necessarily perceived as stressful since their stress levels declined until about 3 am and did not significantly increase until the start of the morning rush hour.

Subjective time pressure ratings depicted in Figure 28 show a similar pattern of results, and are highly intercorrelated with the subjective stress scale results. The range of correlation coefficients, all of which were significant, spanned from r = 0.662 to r = 0.951 with a median correlation value of r = 0.79. The scale value for many of the ratings, especially during first and second shifts, was significantly higher for time pressure than for general feelings of stress. This concurs with the subjective report data previously mentioned in which multitasking between T&E and MOW crews was identified as a stressor. Yet, as these subjective ratings rarely reached even 50, it appears as though these dispatchers, on average, did not report feeling overly stressed.

5.4.2 Objective Stress Level

Salivary cortisol samples were collected every 2 hr while the dispatchers were on shift. Figure 29 is a plot of the amount of cortisol found in the saliva (in units of micrograms per deciliter). The graph shows a well-defined circadian pattern, though there is an additional spike at the start of the evening and night shifts. These data did not show much variation when time-of-day was controlled. Furthermore, these levels were not above the normal physiological level for adults (Laudat, 1988).

Cortisol values within the normal adult population vary little from person to person when time-of-day is controlled. This held true for the current sample (see Figures 29a and 29b). Therefore, the above results are not likely to be due to individual differences. Furthermore, analysis of the cortisol did not show differences related to number of consecutive days of work nor were differences found when examining cortisol levels collected at each interval for the dispatchers working their given shifts. The greatest number of reported consecutive days worked immediately prior to the study's onset was 40, and in fact, many of the dispatchers in this study had worked more than 14 consecutive days prior to the start of the study. Often, these long stretches of continuous work ended during the period when this study was conducted, and all dispatchers took at least one day off during the 14 day study period. This situation provided an unanticipated opportunity to examine the effect of extended periods of work (consecutive days) on dispatcher stress and salivary cortisol.

Comparison of 24 hr cortisol patterns for consecutive days of shift also revealed no significant pattern other than the previously documented circadian variation, with peak cortisol measured at 7 am and lowest cortisol measured between 11 pm to 1 am. This pattern did not vary across consecutive days into shift. Furthermore, there was no indication that those dispatchers who worked up to 40 consecutive days showed any elevation of cortisol level attributable to stress, excessive workload, or excessive consecutive days of work. When

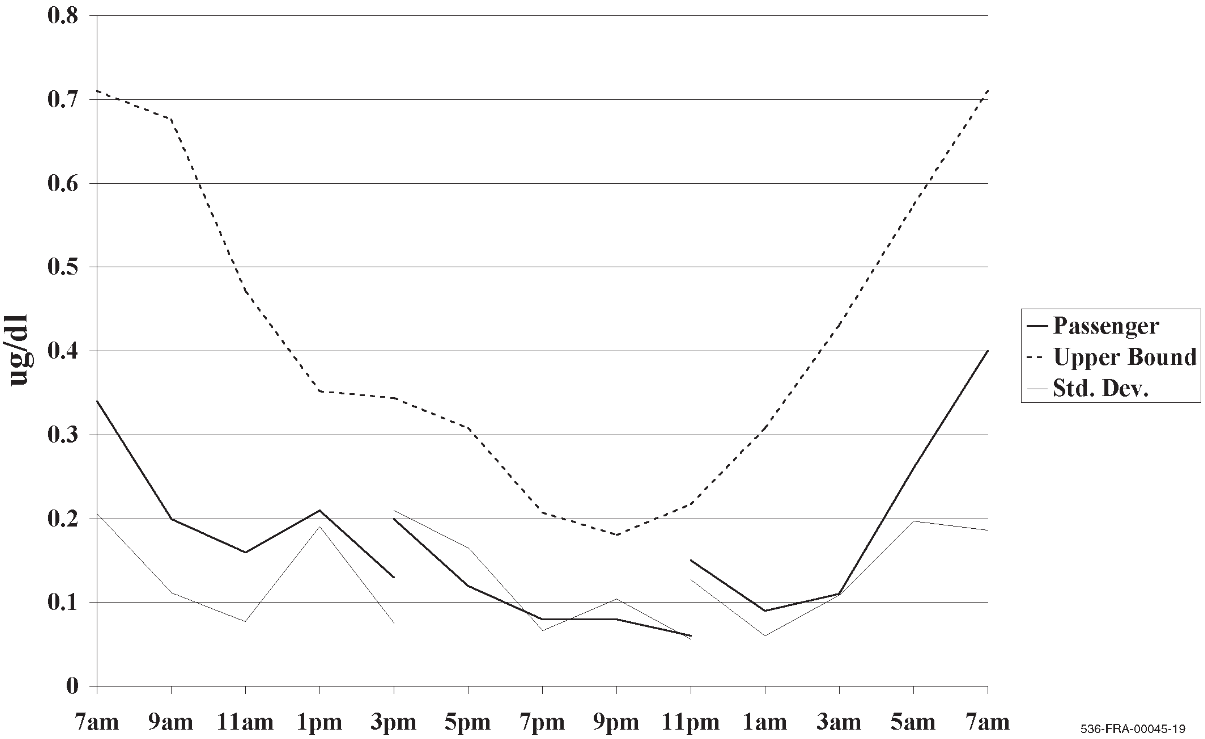


Figure 29a. Mean salivary cortisol levels as a function of time of day for passenger operation by shift

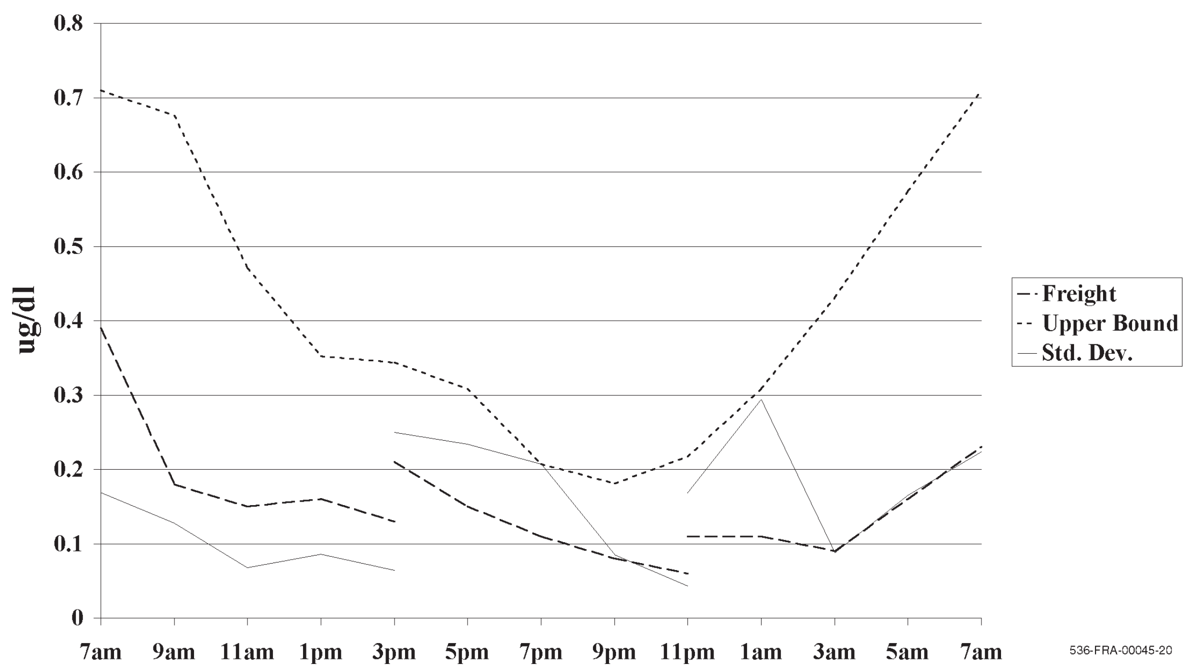


Figure 29b. Mean salivary cortisol levels as a function of time of day for freight operation by shift

controlling for time-of-day, both operations showed a small, though significant negative correlation between cortisol level and shift (freight $r = -0.169$, $p = 0.000$; passenger $r = -0.0785$, $p = 0.36$). This could be interpreted as reflecting a less stressful working environment during the night shift than during the day and evening shifts.

The internal circadian rhythm of dispatchers working regular night shifts or rotating shifts that included night work, as indicated by the cortisol concentration, remained set to a phase position very similar to that seen in human subjects living on a regular night sleep-day activity schedule. The most plausible explanation for this is that these dispatchers work at night in a relatively dark indoor environment, then commute at the end of their shift in bright outdoor morning sunlight. The low activity level required by the job also probably aids in maintaining the circadian system. For physiological purposes, therefore, their light-dark cycles have not been reversed and there is little opportunity for them to reset their circadian phase position to adjust to working night shift.

Participants' cortisol data were also directly compared to normative data. Upper and lower boundaries of normal levels of salivary cortisol were obtained from work published by Laudat and his colleagues (Laudat, 1988). As these data were in nanomoles per liter (nmol/L), they were first converted into micrograms per deciliter, the unit used in this study.[6]

Figures 29a and 29b depict both the averaged 2 hr salivary cortisol values over the course of each shift and the upper bound of normal cortisol values. Dispatcher salivary cortisol levels fell well below this upper bound. Individual analysis of dispatcher cortisol levels also revealed that about 95 percent of the cortisol samples fell below this bound. In fact, only one participant showed consistently elevated levels of salivary cortisol. However, as the figures illustrate, the standard deviation of the cortisol levels was large relative to the mean for the evening and night shift freight dispatchers, indicating a wider range for these groups.

These results were unexpected. A discussion was held with various cortisol researchers, including Dr. Dirk Hellhammer at Trier University, to understand these results. It is possible that these levels were in fact suppressed due to either habituation or burn-out of the cortisol producing system within the brain. Either of these conditions might result if one is placed under high levels of chronic stress. However, Dr. Hellhammer and others similarly concluded that the most likely explanation of the cortisol data, when placed in context with the subjective report and work schedule data, was that the railroad dispatchers were not under a high level of physiological stress.

[6]This conversion was conducted by first multiplying each normative value by the molar mass of cortisol (362.46 g/mol), as found in the 1998 Merck Index, and then dividing by 1 mole to yield normative values in grams per liter (g/L). Each normative value was first converted from nmols/L to mol/L before inputting into the following equation in order to keep mole units of measurement congruent. The conversion formula is as follows: X = normative value mol/L • 362.46 g/mol. Normative g/L values were then converted to µg/dL for comparison with dispatcher cortisol levels.

5.4.3 Summary

There is little physiological and self-rating data indicating that these dispatchers were under a high stress load. This is surprising given the number of items 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 a higher tolerance to stressors or have better coping mechanisms. It was apparent from the case study of the person who left dispatching that this individual felt overwhelmed with job stress and was not able to effectively cope with it. Furthermore, based upon dispatcher self-reports of their sources of stress, the workload itself seems to be only one of many possible components in what appears to be a multivariate construct. The balance of these various sources, and inherent individual differences makes it practically impossible to identify root or primary sources based on the current sample size.

It is also likely that, as with workload, the cortisol and on-duty self-report data were collected too infrequently. Salivary cortisol data also has the added problems of possessing a long delay between the stress trigger and cortisol release, and its ease of being masked. Collecting blood cortisol levels or increasing the sampling rate was not possible given the in situ constraints. In addition, no performance data exists to compare with the various stress levels. It may be that the level of reported perceived stress does lead to changes in performance even if the entire length of the 100 mm scale was not utilized. Based on the available data, however, it appears that these dispatchers are not experiencing excessive stress.

5.5 Fatigue

Railroad dispatcher fatigue was explored through both objective and self-report measures. First, sleep-wake behavior was quantified using Actiwatch data. Next, background survey items asked specific questions related to participants' sleep, sleepiness and fatigue. Third, participants rated their level of subjective fatigue every 2 hr while on duty using a 100 mm VAS. Lastly each dispatcher completed a daily sleep log after their primary sleep period. The sleep log asked questions pertaining to the timing and latency of sleep periods, whether sleeping aids were taken, timing of naps, and various sleep quality and sleepiness scales. The sleep logs were intended to complement the data collected from the Actiwatches.

Based upon the literature and anecdotal information, the dispatchers were expected to show increased subjective fatigue both throughout their work shift and across the workweek. Workday/nonworkday differences in sleep length and quality were also expected, especially for those working nights and irregular shifts. It was also expected that the night shift workers would take more naps than either the day or evening shift workers.

5.5.1 Subjective Fatigue

Subjective fatigue data were collected using at-work self-report forms. As a starting point, all subjective self-rating measures recorded at the start of the shift, including subjective fatigue, were compared to those recorded at the end of the shift. All six subjective measures showed a significant worsening between the start of the shift and the end of the shift (see Table 27). Table 28 partials out these significant effects by shift. All six subjective ratings showed significant increases from beginning to end of the day shift. There were fewer differences found during the evening and night shifts. Specifically, fatigue and stress showed significant increases from beginning to end of the evening shift, while only fatigue showed a significant increase from the beginning to end of the night shift.

All six subjective ratings were also compared between study sites. Very few differences were found between these two groups with respect to the subjective workload, stress and fatigue

Table 27. Comparison of the workload, stress and fatigue scales (100 mm VAS) filled in at the beginning and end of work shift by all participants on all shifts

Variable	*t* Value	df	*p*	Start Value	End Value	% Difference
Mental Demands	-3.476	344	0.001	44.31	49.43	5.12
Time Pressure	-4.173	344	0.000	40.57	46.62	6.05
Level of Effort	-2.925	344	0.004	46.41	50.54	4.13
Overall Workload	-2.542	344	0.011	40.43	44.15	3.72
Fatigue	-14.025	344	0.000	21.39	40.12	18.73
Level of Stress	-7.907	344	0.000	26.63	37.96	11.33

Table 28. Comparison of the workload, stress and fatigue scales (100 mm VAS) grouped by shift, filled in at the beginning and end of work shift by all participants

Shift	Variable	*t* Value	df	*p*	Start Value	End Value	% Difference
1	Mental Demands	-5.735	151	0.000	35.98	49.25	13.27
1	Time Pressure	-5.911	151	0.000	32.43	46.19	13.76
1	Level of Effort	-4.806	151	0.000	39.78	50.45	10.67
1	Overall Workload	-4.186	151	0.000	34.13	44.51	10.38
1	Fatigue	-10.620	151	0.000	19.66	40.13	20.47
1	Level of Stress	-7.841	151	0.000	21.95	39.92	17.97
2	Demand	0.609	97	0.544	32.76	31.46	3.97
2	Time	0.032	97	0.974	31.27	32.43	3.58
2	Effort	0.307	97	0.760	32.59	33.49	2.69
2	Workload	0.042	97	0.966	24.96	27.37	8.81
2	Fatigue	-6.136	96	0.000	17.02	33.23	16.21
2	Level of Stress	-3.047	97	0.003	24.88	32.94	8.06
3	Demand	0.404	94	0.687	26.95	28.89	6.72
3	Time	-0.023	94	0.982	25.00	27.68	9.68
3	Effort	0.530	94	0.597	24.67	27.50	10.29
3	Workload	1.317	94	0.191	23.82	24.92	4.41
3	Fatigue	-7.198	94	0.000	28.62	47.14	18.52
3	Stress	-1.840	94	0.069	22.16	24.35	8.99

ratings at the beginning and end of shift. As can be seen in Table 29, the dispatchers from the passenger operation reported higher levels of mental demand at the start of shift and higher levels of mental demand and fatigue at the end of shift. No other significant differences were found.

Next, subjective fatigue reports over the course of the shift were compared between the two field sites. Figure 30 displays the mean reported levels of subjective fatigue by time-into-shift and dispatching site. There were significant differences between freight and passenger operation dispatchers at 9 am, 11 am, and 1 pm. No other statistically significant between-site differences existed for any of the shifts. As mentioned above, levels of fatigue significantly increased throughout the duration of the shift, doubling or nearly doubling in all cases. Interestingly, no significant dip in fatigue level was detected mid-shift, even though increased fatigue levels are often found to occur after lunch, especially during the day shift. Fatigue ratings for the start and

Table 29. Comparison of the workload, stress and fatigue scales (100 mm VAS) grouped by type of operation, filled in at the beginning and end of work shift by all participants on all shifts

Variable	*t* Value	df	*p*	Freight	Passenger	% Difference
Mental Demands, Start of Shift	-2.082	352	0.038	41.10	47.78	6.68
Fatigue, Start of Shift	-3.232	351	0.001	18.33	25.32	6.99
Fatigue, End of Shift	-2.522	351	0.012	36.47	43.62	7.15

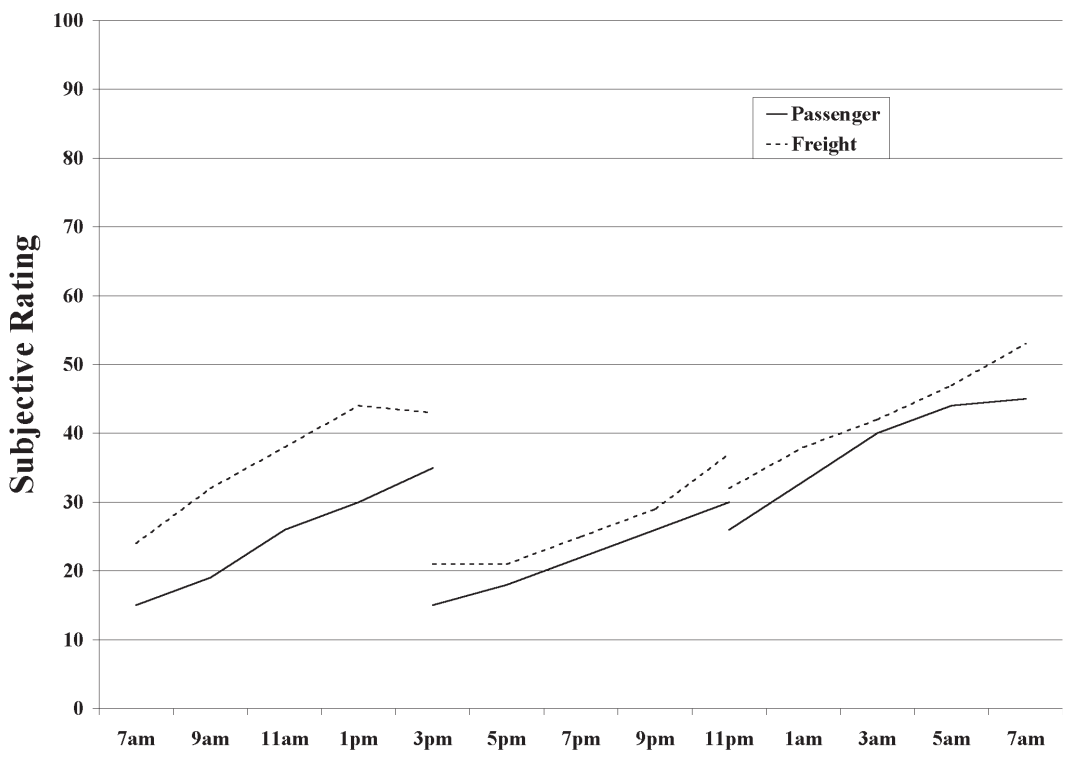


Figure 30. Subjective fatigue ratings as a function of time of day for passenger and freight operations

end of the night shift were significantly higher than for the day and evening shifts. This is likely to be due to the dispatchers fighting their circadian rhythm as well as handling the responsibilities of being on duty.

5.5.2 Primary Sleep

The dispatchers were asked a series of questions pertaining to the quality of their sleep. Working shifts, especially night and irregular shifts, has been shown to negatively impact the duration and quality of one's sleep, putting individuals at risk for acute and chronic sleep deprivation effects (Tepas, Paley, and Popkin, 1997; Tepas and Carvalhais, 1990; Tepas, Walsh, Moss and Armstrong, 1981; Johnson and Naitoh, 1974).

Table 30 presents rate and frequency data on dispatcher problems of falling asleep, frequent awakenings, and early awakenings, as well as their rate of waking up tired on workdays and nonworkdays. These data are presented both by shift and aggregated.

About one-third of the dispatchers reported having "problems falling asleep" at least twice a week. The evening shift dispatchers reporting problems falling asleep twice as often as the other two shifts. This was surprising given that, according to the literature, those working the night shift are usually the ones with the highest incidents of sleep-related difficulties. The rationale most frequently cited for this finding is that people working this shift are generally less free to sleep when they want and at an appropriate physiological time (Tepas and Monk, 1987). The observed difference in this study may be due to some shift by lifestyle interaction unique to this sample or population. Problems with frequent awakenings seem more consistent with the literature, with people on the night shift and extraboard showing a significantly higher rate. Interestingly, of those dispatchers who reported frequent awakenings, the frequent awakenings seemed to have occurred between four and five days a week, regardless of shift.

With the exception of the extraboard dispatchers, similar rates were found with problems of "early awakenings." Again, it was expected that the night shift workers who slept days would

Table 30. Rate and frequency of sleep-related problems

	Shift				
	Day	Evening	Night	Extra Board	Overall
Problem Falling Asleep (% yes)	33	33	40	27	32
Frequency of Problem (Days/Week)	2.5	4	2	2	2
Frequent Awakenings (% yes)	33	33	60	46	46
Frequency of Problem	5.5	4	5	2	4
Early Awakenings (% yes)	33	33	60	27	35
Frequency of Problem	5.5	2	3	3	3
Waking up Tired (% yes):					
Workdays	33	17	80	36	41
Nonworkdays	33	17	40	9	22

have problems staying asleep given their circadian system, a system which showed no sign of adaptation based on salivary cortisol (see subsection 3.7.1). Those on the extraboard may have a lower rate of occurrence given that: 1) they are not always sleeping days since they may also work the day and evening shift, and 2) as mentioned, it does not appear that any of the participants' circadian rhythms were significantly altered from a typical diurnal rhythm.

The rate of those "waking up tired" on workdays and nonworkdays shows a pattern consistent with the literature. Specifically, those working nights have the highest rate of this symptom, and this problem is dramatically reduced when they return to nighttime sleep. Again this is probably due to a lack of adjustment in the dispatchers' circadian rhythms. The reduction in the rate between workday and nonworkday waking-up-tired self-reports of the extraboard dispatchers is also not surprising. This apparent effect could not be explored or confirmed with the current data due to the small sample size and the variability in the dispatchers' schedules.

The sleep log data was grouped by whether the dispatcher went to work, a workday, or was off, a nonworkday. The following sleep quality measures were examined: ease of falling asleep, ease of arising, sleep duration, overall sleep quality, the Karolinska Sleep Scale and Naval Psychiatric Research Unit (NPRU) positive and negative mood state scores (see Figures 31a, 31b and Figure 32). The first five scales are discrete in nature, while the NPRU approaches a continuous scale. In addition, the number of naps taken during each day and the sleep length during dispatchers' primary sleep periods, as reported on the sleep logs, were also examined.

The first set of analyses tested the above-mentioned sleep quality data separately for the freight and passenger operation dispatching sites. The only difference found was for the freight operation. Specifically, it was slightly easier (1.57 nonworkday, 1.81 workday, where a lower scale value here denotes greater ease of falling asleep) to fall asleep on nonworkdays ($t = -2.236$, df = 260, $p = 0.026$).

Analyses were also conducted to determine the differences between the freight and passenger sites separately for the workday and nonworkday groupings. Only three differences were found, two during nonworkdays and one during workdays. As seen in Table 31, the dispatchers at the passenger operation had a slightly, though significantly lower (as signified with a higher scale value) overall quality of sleep both during their workdays and nonworkdays. The dispatchers at the freight operation, however, found it slightly more difficult to arise after their main sleep period on nonworkdays.

The data presented in Figures 31a and 31b indicate a relatively high level of satisfaction with sleep quality, independent of shift worked or whether the sleep occurred during a workday. Generally, those dispatchers working the second shift reported being more satisfied with their sleep quality than the first shift who, in turn, were more satisfied than the third shift. This trend is supported in the shiftwork literature (Gordon, Tepas, Stock and Walsh, 1979). Significant differences were not found between the first and third shifts nor were significant differences found between workdays and nonworkdays. The NPRU scores, plotted in Figure 32, also show very little variation among shifts and between workdays and nonworkdays. The positive scale ranges between 0 and 57 and the negative scale between 0 and 30. The mood scores are used as

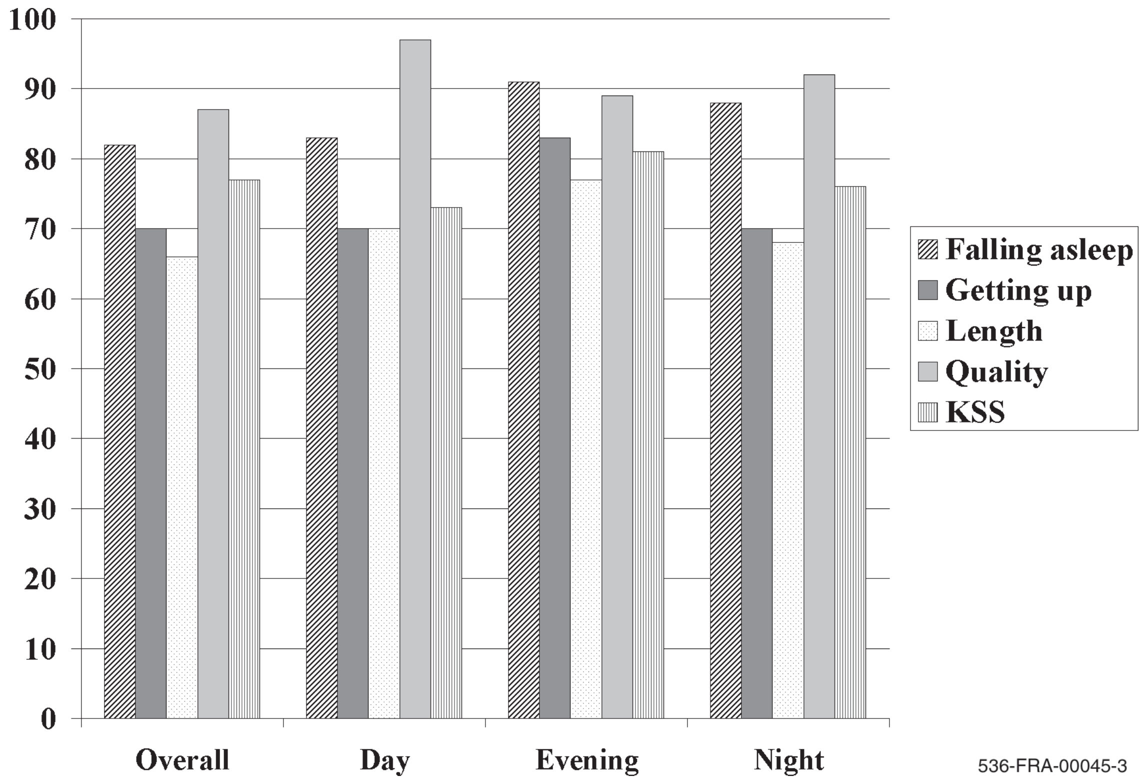


Figure 31a. Measures of sleep quality by shift and overall

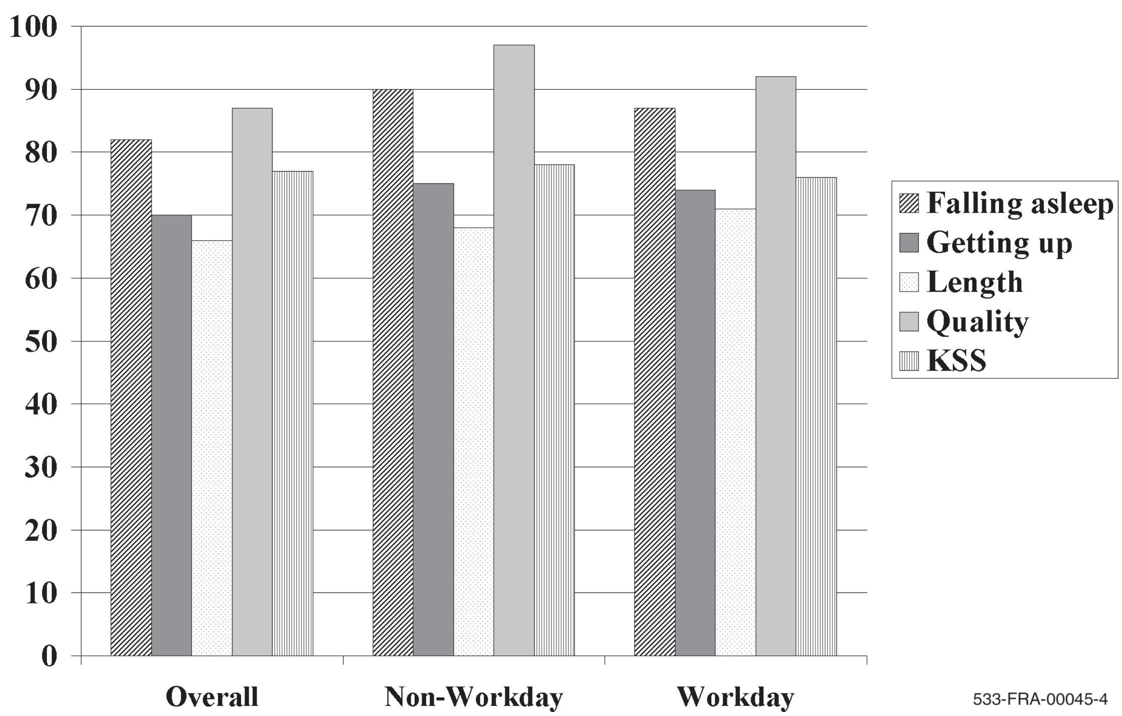


Figure 31b. Measures of sleep quality by type of day and overall

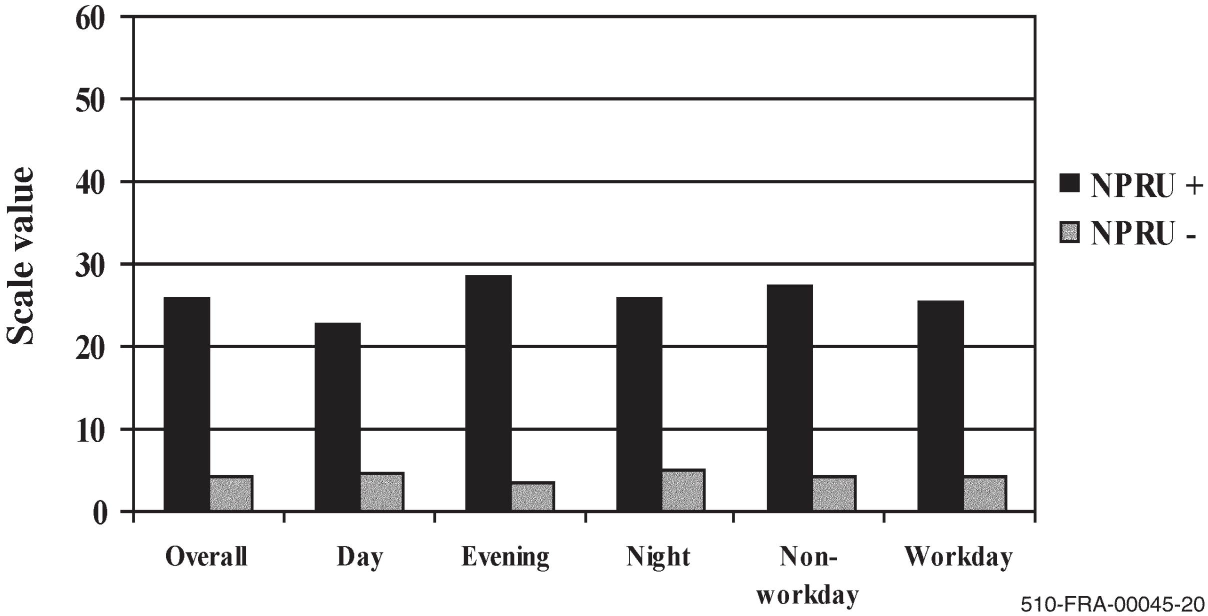


Figure 32. NPRU mood scales

Table 31. Comparison of the sleep quality scales grouped by type of operation and workday/nonworkday, filled in after each daily main sleep period

Variable	*t* Value	df	*p*	Freight	Passenger
Sleep Quality - Workday	-4.059	349	0.000	2.12	2.48
Sleep Quality - Nonworkday	-2.637	127	0.009	1.93	2.28
Ease of Arising - Nonworkday	2.240	127	0.027	2.32	2.06

indicators of potential chronic sleep deprivation. As these values were mostly correlated with sleep length at r=0.3, and not statistically significant, there does not appear to be any indication of chronic sleep deprivation from these dispatchers.

Sleep quality was also measured using sleep efficiency and sleep fragmentation measures that were computed by the Actiwatch software. Sleep efficiency refers to the percentage of time in bed spent sleeping while sleep fragmentation refers to the amount of prolonged movement during sleep, suggesting Stage 1 sleep or wakefulness. These data are presented in Figure 33. In general, sleep fragmentation was around 30 percent; in other words, 70 percent of the sleep was continuous and undisturbed with awakenings. This result did not significantly change when examining the data by shift or by whether the sleep occurred on a workday or a nonworkday.

Figure 34 provides four different measures of sleep length data collected from the dispatchers. These data are presented by shift (i.e., day, evening and night) and by workday/ nonworkday. Subjective sleep refers to how much sleep the dispatcher reported he or she received, while the Actiwatch sleep length measure computes start and end sleep times based on dispatchers' restfulness. Intermittent awakenings that occurred during the sleep time were

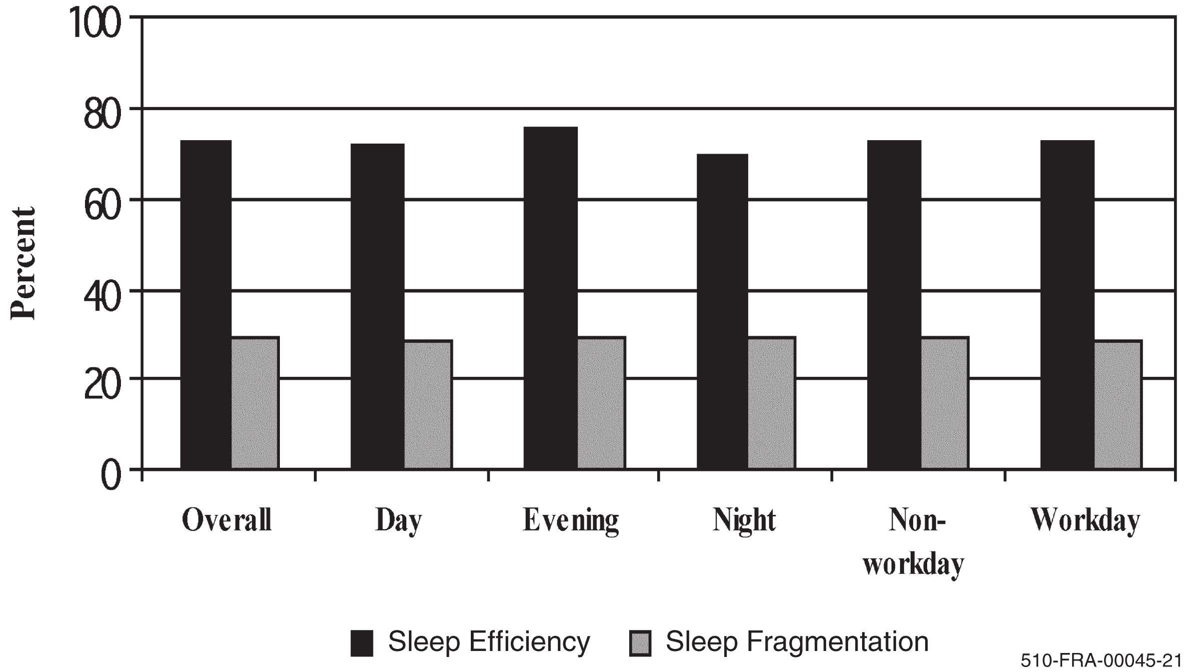


Figure 33. Sleep quality via Actiwatch

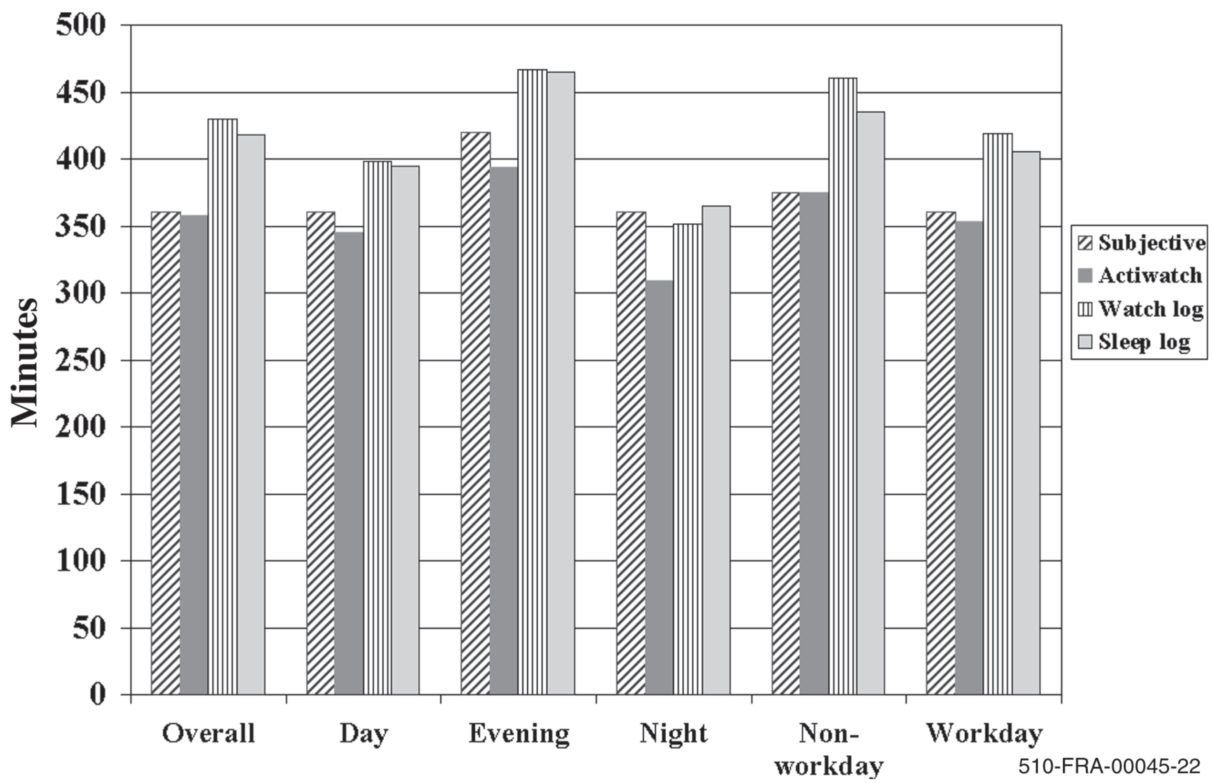


Figure 34. Average sleep lengths using different data collection methods

subtracted out of the sleep length. The other two measures reported here, actiwatch log and paper sleep logs, record sleep start and end times without considering sleep latency or micro-awakenings. Results are quite different between the first two views and the second two views. When considering the first two views, subjective ratings and Actiwatch data, these dispatchers, on average, received between 6 and 7 hr of sleep per day. Dispatchers on evening shifts received the most sleep followed by day and night shift workers. There is no significant difference between workday and nonworkday main sleep durations. These results are amplified when using the second two sleep length measures, and they show a greater workday/nonworkday differentiation. Based on these data dispatchers received between seven and 8 hr of sleep per day, regardless of workday/nonworkday status. Interestingly, these measures were neither highly nor significantly correlated with the sleep quality measures. An additional analysis revealed that those dispatchers who had five or more years of experience slept 1 hr less, on average, than those with less experience. This may, of course, be merely a function of age, though the median age of the sampled workforce was under 40 years.

When asked whether they napped or took sleep aids, napping was found to be the more prevalent strategy (see Figure 35). Those working the evening shift reported napping the least, less than 30 percent, while those on the night shift napped the most, more than 60 percent. This was expected. What was unexpected, however, was that the night shift workers took significantly fewer sleep aids than either day or evening shift dispatchers. There were also no significant differences between workday and nonworkday responses. Table 32 presents the rate and frequency of use of split sleeps and napping. A nap was defined as a sleep period of less than 3 hr duration while a split sleep was defined as two discrete sleep periods with more than

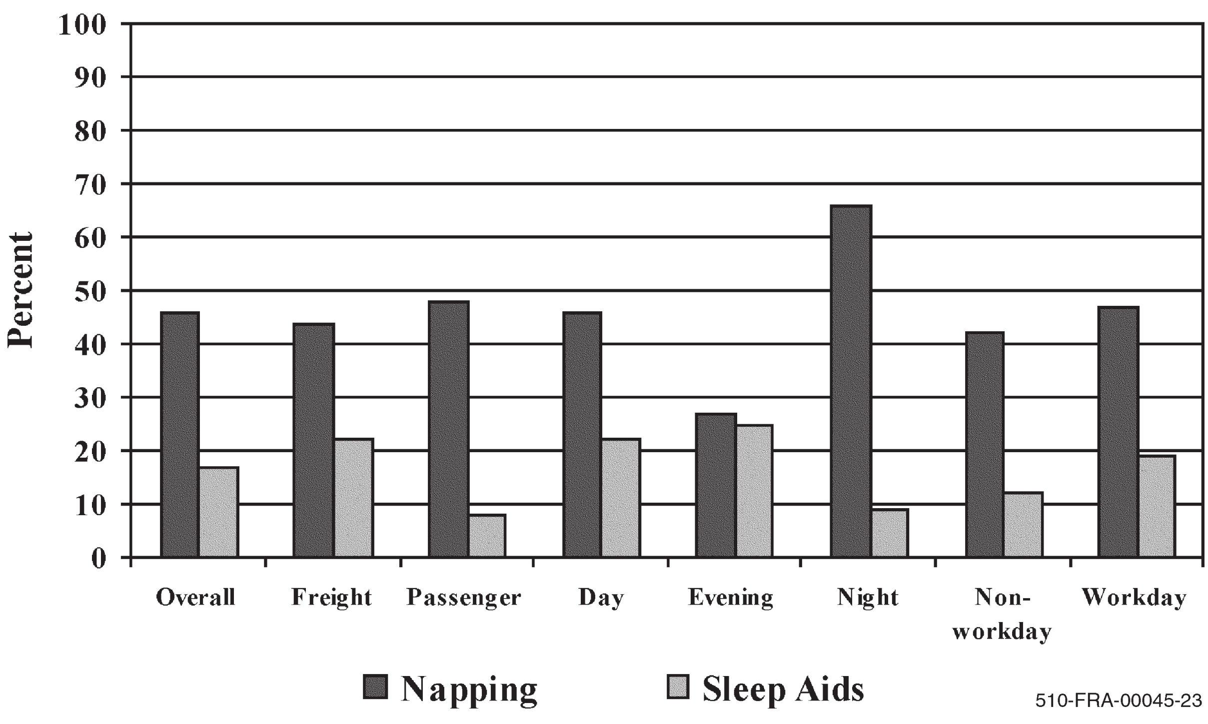


Figure 35. Coping strategies for adequate sleep

Table 32. Usage of split sleeps and napping

	Shift				
	Day	Evening	Night	Extra Board	Overall
Split Sleep (% yes)	17	33	100	72	51
Number Split Sleeps/Week	2	1	6	2	1
Number of Naps per Week	3.5	2.5	5	2	3
Nap Length (min)	30	42.5	75	35	37.5

1 hr of wakefulness between the periods. Again it is the dispatchers working on the night shift who report the highest rate of split sleeps and the frequency of their use. They also report the highest rate of napping and for the longest durations. Figure 36 combines the mean sleep and nap lengths for the various shifts and workday/nonworkday status. This provides a complete view of the amount of sleep a dispatcher typically receives during a "day." Napping, it seems, allows the night shift workers to obtain roughly the same amount of sleep as day workers. The evening shift workers, however, still sleep significantly longer, overall. A significant difference was also found between workdays and nonworkdays.

5.5.3 Nap Patterns

The nap data was analyzed separately in order to better describe and evaluate its duration and efficiency. These data were also used to test the compatibility of the Sleep Log and Actiwatch methodologies. Regarding nap lengths, it appears as though those working evening and night shift nap significantly longer than those on day shift but no significant difference was found

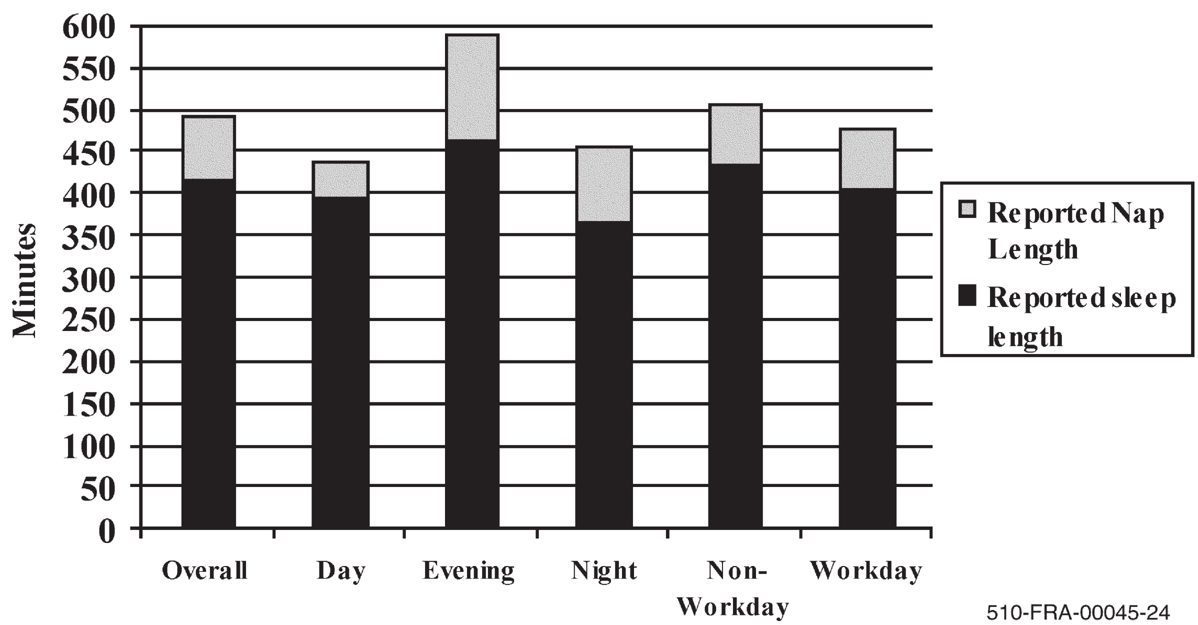


Figure 36. Daily reported sleep and nap patterns

when comparing the evening to the night shift (see Figure 37). Similar results were found for both Sleep Log and Actiwatch data. It was surprising that those nappers working evening shift had longer nap lengths than those working morning shift. The literature suggests that the evening shift has the greatest opportunity to sleep as long as they wish during their main sleep period. Furthermore, background survey data revealed only minimal moonlighting, and those were either home businesses or church/organizational work. It may be that these dispatchers either split their sleep intentionally or engaged in recreational napping. The present data is inadequate to explain the observed napping pattern. There was no statistical difference between workday and nonworkday nap lengths indicating that perhaps naps are more a part of these dispatchers' lifestyle than a means for reclaiming any lost sleep.

Some significant differences were found when comparing the nap data collected from the Actiwatch and from the Sleep Log. These differences were found when considering just the overall values and for both workday and nonworkday values. Differences were not found within the individual shifts, though this is likely due to low power. In all cases, averaged self-reported nap lengths recorded on the sleep logs were longer than those recorded by the Actiwatches. The most likely explanation for the shorter Actiwatch nap durations is that, while the dispatchers recorded when they started to nap and when they stopped, the Actiwatch subtracts sleep latency (i.e., restlessness at the beginning of a sleep or nap period) from the nap length, considering the individual to be in a wakeful state.

To explore the *quality* of the naps, the percent of nap efficiency and percent of nap fragmentation data were graphed and analyzed (see Figure 38). Nap efficiency refers to the amount of quiet sleep time (i.e., slow wave sleep) over a given sleep period, while nap fragmentation refers to the amount of movement recorded during the sleep period. Both measures are recorded as percent values. The amount of fragmentation is relatively low,

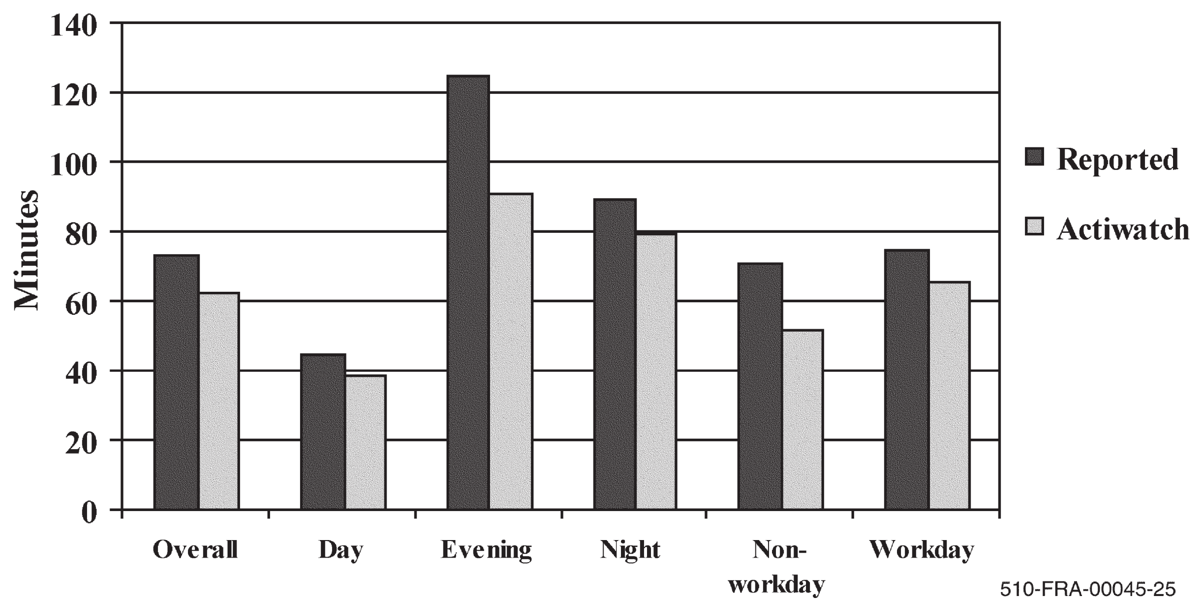


Figure 37. Average nap lengths

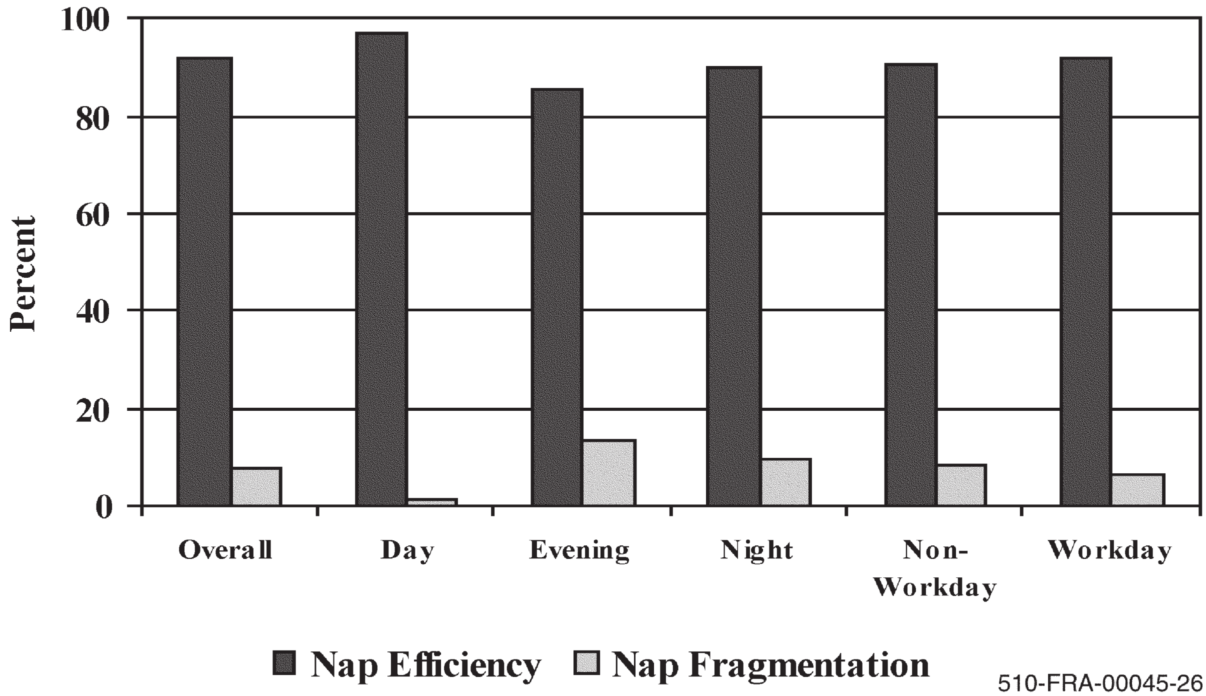


Figure 38. Nap quality via Actiwatch

10 percent or less in all but the dispatchers working the evening shift. The increased nap duration for the evening shift workers likely compensates for the increased sleep fragmentation. In fact a significant Pearson correlation of r=0.88, p<0.05 was calculated between these variables. No other significant correlations were found for the other two shifts. Perhaps temporal placement of the nap is a better explanation for the differences in fragmentation. Specifically, those working the evening and night shift were more apt to sleep and nap in the morning hours, conflicting with the circadian system. Sleeping at physiologically inappropriate times due to the circadian system often leads to less efficient sleep, and therefore a longer sleep duration may be required to obtain a sufficient amount of sleep. Those working the day shift only had the late afternoon and evening to nap, times that lead to more efficient sleep according to circadian theory (Wever, 1979).

Nap lengths were then separately added to sleep lengths calculated from the sleep logs (Figure 36) and from the Actiwatches (Figure 39). When the primary sleep periods and naps were combined, the average daily sleep times reported by the sleep logs for those who napped rose from about 410 min to 460 min, almost 8 hr. No significant differences were found between the different shifts, although the third shift workers who napped did, on average, sleep 50 min less than the first and second shift workers. A significant difference was found, however, between average workday and nonworkday total sleep length for nappers. Specifically, the total sleep length on nonworkdays was significantly longer than on workdays. Given what was found regarding nap length, this would suggest that while napping behavior does not seem dependent upon whether one is working, these dispatchers take advantage of the off days to extend their main sleep period. This is consistent with the current literature.

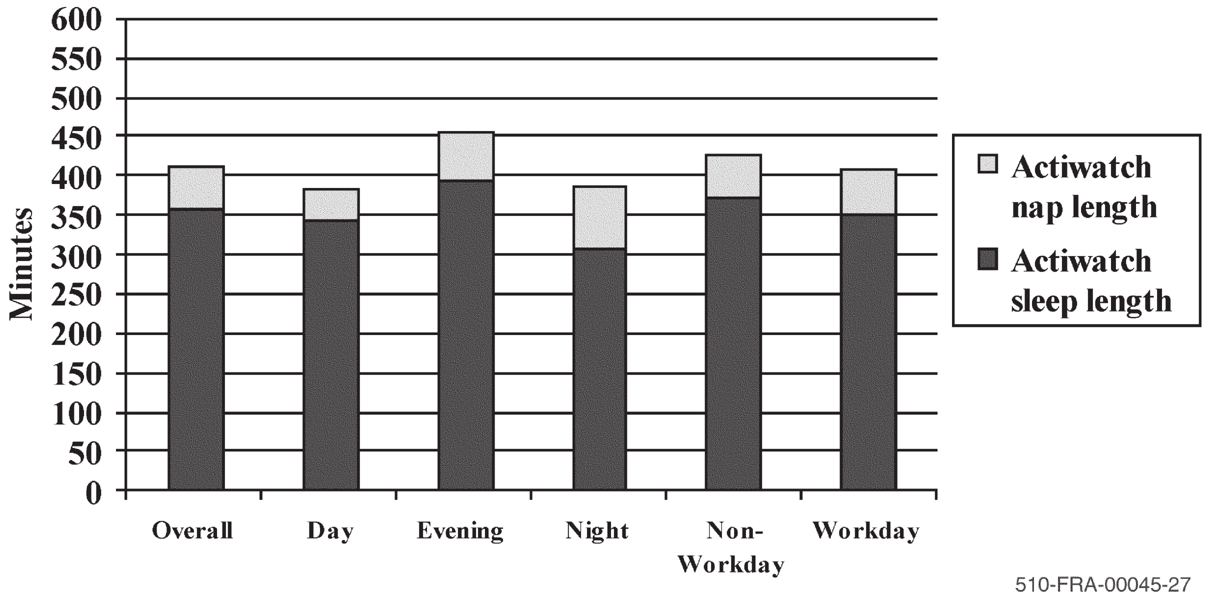


Figure 39. Daily Actiwatch sleep and nap patterns

The sleep data gathered by the Actiwatch provided somewhat different results. This was expected as this sleep length data did not include sleep latency duration. When nap duration is included, overall sleep length increased from 355 min to about 380 min, well under the 460 min indicated by the sleep logs. Unlike the sleep log data, significant shift differences were found with respect to total sleep length while the workday/nonworkday difference disappeared. The evening shift dispatchers who napped had a significantly longer total sleep time than either the day or night shift dispatchers. No other differences were significant. Again, this indicates that the napping strategy, like sleep, is affected by the time at which it is taken, and therefore, provides varying amounts of additional sleep. This is also in accordance with the sleep fragmentation results whereby day naps are usually taken by night and afternoon shiftworkers, and therefore would be most susceptible to circadian disruption.

5.5.4 Summary

The sleep patterns of this study population appear to be typical of other shiftworkers. Specifically, those on the afternoon shift had the longest sleep lengths while those on night shift slept the least. Napping appears to be the most utilized strategy for gaining additional sleep. This extra sleep, when combined with the primary sleep, increased the overall sleep time of the night shiftworkers as to equate them with the day shiftworkers. The afternoon shiftworkers still obtained significantly more sleep than did the other two shifts.

There is little evidence to support anecdotal information of severe, acute sleep debt. Furthermore, the only indication of any chronic sleep debt or its resulting increase in fatigue in this study population was lower than expected NPRU positive scale scores. However, NPRU negative mood scale scores have been shown to be more predictive of such fatigue states than the positive scale scores, and they were, on average, quite low for all participants. Furthermore, the subjective fatigue scale and the self-report sleep log items did not indicate any extreme or strong tendency towards feelings of sleepiness or fatigue while awake.

It was shown, though, that feelings of fatigue did increase monotonically throughout the course of a dispatcher's shift. While these fatigue ratings rarely increased above 60 for an 8 hr shift, shifts of a longer duration, or double shifts given emergency staffing situations, could result in substantially more fatigue. A rest break may ease both the level and rate of increase in perceived fatigue. This hypothesis, however, remains to be tested.

Regarding the future use of the sleep log and Actiwatch methodologies, both can provide meaningful data. The sleep log is a low cost tool that seems to be reliable with the current subject population. It is especially useful for considering the placement and number of sleep periods and naps, and to gain an approximation of sleep duration. However, if specific amounts and quality of sleep are sought, then actigraphy data, in this case from an Actiwatch AW4-32, in conjunction with a sleep log, may be a more appropriate methodology. It is currently unknown, though, how well the Actiwatch sleep data, determined from the user's activity level, parallels data collected from EEG recordings for these participants who spent a considerable amount of time quietly monitoring their territory. While the product itself has been validated using EEG recordings, it is impossible to say with certainty how well it works in this type of field setting (McConnell, 1998).

6. KEY FINDINGS AND RECOMMENDATIONS

This field study offers some important insights into the job demands of a railroad dispatcher and the related levels of workload, stress and fatigue. In reviewing the findings, it is important to keep in mind that the study involved only 37 dispatchers from two dispatching centers for a two-week data collection period. Therefore, the results may not be generalizable to other centers or to the nation's dispatcher population as a whole. This section presents key study findings with respect to dispatcher workload, stress and fatigue as well as lessons learned regarding the field methodology for future studies of these factors. The section concludes with some recommendations for future research regarding the railroad dispatching environment and related workload, stress and fatigue of dispatchers.

6.1 Key Study Findings

The study findings are presented below to address each of the objectives of the study.

6.1.1 Dispatcher Health

Comparison of the results of the background survey with normative epidemiological data indicates that participants in the 25 to 44 age group reported incidences of back pain, headaches and skin disorders that significantly exceeded adult norms for this age group. There are many possible explanations for these problems. Back pain could be due to poor seated posture resulting from a chair not suited to the individual. Headaches can result from poor lighting in the work area, an uncorrected vision problem or glare from a computer screen. Skin disorders can have any number of causes. All three problems can also be the result of stress. This finding suggests a further investigation into health-related problems in the U.S. dispatcher population. In addition, dispatching center management might want to determine the extent to which any aspects of their work environment could contribute to these disorders at their center.

6.1.2 Workload

Available activity count data, in terms of number of Form Ds, route blocks, and number of train and other track users, appear to be insufficient to gauge the level of the dispatcher's workload. These factors accounted for less than half of the variance in the subjective workload variations. This suggests that dispatcher workload is a more complex construct than anticipated. Future measurement of dispatcher workload should either take into account additional dispatcher activities or determine more realistic cognitive loads for these activities. Given the planning and decision-making duties of a dispatcher, future measurement of a dispatcher's workload should also take into account these cognitive aspects of the job.

Though the subjective workload ratings and work volume counts did not appear to capture the variability of the dispatcher's workload, the mTAWL method succeeded in this area. The mTAWL method proved valuable for identifying the cognitive, visual, auditory and psychomotor components of workload for a given desk and for documenting them over time. Due to its labor-intensive nature, the mTAWL may not be an ideal data collection tool for a research study, but its potential as a diagnostic tool is significant. The mTAWL appears to be very appropriate for identifying and documenting workload at a particular dispatching desk, and for comparing workload over time, or across desks within a dispatching center. Ultimately, the value and utility of the mTAWL technique for researchers and operations personnel must be weighed against its labor-intense methodology.

6.1.3 Stress

Participants' responses to the background and exit surveys indicated some level of workplace stress. Forty-three percent of the participants reported "a lot" or "a great deal" of anxiety and a third reported calling in sick due to stress at least "sometimes." Nearly all participants (92 percent) reported that other dispatchers "sometimes" or "frequently" lose their temper at work. However, the sources of this stress varied by site and by level of experience. Some of the stress was based on the dispatcher's job demands, but a significant amount was also attributed to other aspects of the work environment. This suggests that stress in these dispatchers is multivariate in nature and not completely centered on the work itself.

In contrast to the results of the surveys at both sites, there was little physiological or self-report data from the field study indicative of a high level of stress in the study participants. However, this should not be interpreted as an indication that workplace stress does not exist. Since the dispatcher's workload and associated stress can change rapidly over short periods of time due to the inherent nature of the job, it is more likely that the collection of subjective ratings and the measurement of cortisol, a hormone easily masked by other events, every 2 hr was inadequate to capture variations in the dispatcher's stress level. Two other factors may have affected stress levels. First, as data was collected in late summer, it is entirely possible that these dispatchers either had recently returned from vacation or were looking forward to an upcoming one. In addition, seasonal variations in traffic could be responsible for the low workload and stress observed in this study.

6.1.4 Fatigue

Both physiological data and self-report information indicate that dispatchers' sleep patterns are typical of shiftworkers. Those on the afternoon shift had the longest sleep lengths while those on the night shift slept the least. Those working the night shift also relied the most on napping to gain additional sleep. While the study produced little evidence to support either an acute or chronic sleep debt for this group, study participants did report an increase in perceived fatigue over the course of the shift. The use of planned breaks during the shift is one recognized strategy for alleviating this feeling of fatigue. As was the case with the subjective ratings for workload and stress, ratings of perceived fatigue every 2 hr were probably inadequate to capture moment-to-moment variations.

6.1.5 Methods

This initial dispatcher study has provided important insight into possible modifications and enhancements to both the measures and protocol for future studies of dispatcher workload, stress and fatigue. As the above findings indicate, the measures of dispatcher workload used in this study may not capture the cognitive or mental components of the dispatcher's job. This is because so much of a dispatcher's work is "in the head" with little or no visible evidence of work. Future workload measures and protocols should try to incorporate these aspects of a dispatcher's work. Ultimately, it will also be necessary to relate any measure of dispatcher workload to dispatcher performance.

With respect to a physiological measure of stress, recent research has shown salivary amylase or salivary melatonin to be a more responsive marker of stress than salivary cortisol, and to be much more resistant to masking effects. In addition, ambulatory stress monitoring techniques such as electromyogram (EMG), electrocardiogram (EKG), galvanic skin response (GSR), or pupillary data should also be considered because they provide continuous data.

Regardless of the workload, stress, and fatigue measures used to collect data during a dispatcher's work, more frequent data collection is necessary to capture the variation, and short-term fluctuations in workload, that are inherent in the dispatcher's job. Data collected every 2 hr may be insensitive to these variations. However, based on feedback from dispatchers solicited through a focus group during the pilot study, increasing the frequency of data collection in the workplace will probably not be acceptable to either railroad management or dispatchers. Because these individuals are conducting safety-critical work, there is also an increased risk of distracting them from their work as more is asked of them. A medium- to high-fidelity dispatching simulator would be more suitable for this type of research. Specifically, a dispatching simulator would enable researchers to control workload and other conditions of the work environment in order to see their effects on the dispatcher. A dispatcher simulator would permit collection of subjective and physiological data as frequently as needed and dispatcher performance data could be easily collected.

The *in situ* nature of this study limited the analyses that could be performed in two important respects. First, it was not possible to control any aspects of the work environment or the work itself, e.g., workload. Second, given the constraints of the workplace there were limits on how much data could realistically be collected given that the dispatcher had to perform his/her job. These two limitations made it difficult to study individual workload, stress and fatigue factors in depth. The small study population also limited examination of the interactions among these three variables. Regardless of the method and measures used to study railroad dispatcher workload, stress and fatigue, each of these topics should be explored separately, before interactions among them are examined. In doing so, researchers would be able to study each topic in more depth than was possible in the current study, leading to a more complete understanding of the causes and underpinnings of dispatcher workload, stress and fatigue.

6.2 Recommendations for Additional Research

The field study described in this report was the first to explore railroad dispatcher workload, stress and fatigue concurrently. While the present study gave insight on some aspects of the dispatcher's job and work environment, many issues remain to be explored. The findings and experience of this study suggest several issues for further research. These issues include the following:

- *Dispatcher Performance Measures* – The current study did not include any measures of dispatcher performance. Without performance measures, it is not possible to say how the dispatcher's ability to adequately perform his duties is affected by increasing levels of workload, stress and/or fatigue. A field or simulator study that includes dispatcher performance measures can help determine how increasing levels of one or all three factors compromise the dispatcher's ability to carry out the job responsibilities in an efficient and safe manner. These performance measures should include both safety-related and efficiency-related outcomes.

- *Comprehensive Dispatcher Profile* – The small number of participants in this study and the fact that they were from only two dispatching centers make it difficult to draw conclusions about the U.S. dispatcher population as a whole. Administering a survey to all U.S. dispatchers would lead to a much more comprehensive profile of the U.S. dispatcher workforce. This profile would help to identify any population-wide characteristics that may pose safety risks to railroad operations.

- *Rest Breaks* – Although study participants reported taking about four breaks per shift, two-thirds reported difficulty getting coverage so that they could take the break. Study results also indicated increasing fatigue throughout the workday across all three shifts. Planned rest breaks are a recognized strategy for reducing fatigue. To date only one dispatching center in the U.S. has implemented planned rest breaks and this was only for a limited time. Another center experimented with having an extra dispatcher available to cover during breaks but currently no centers have scheduled breaks. A field or simulator study could explore the value of planned rest breaks for reducing stress and fatigue in the dispatching environment.

- *Prediction of Loss of Alertness* – Both the Army and Air Force have been working on models for predicting performance degradation due to fatigue. Other models also exist. Development of a loss of alertness model for dispatchers would give both the dispatchers and their management a tool for assessing when performance is likely to degrade due to fatigue.

- *Measures and Models of Dispatcher Workload* – There is a need for valid and reliable measures of dispatcher workload. These measures should account for the duration and intensity of the workload over time, and individual differences among dispatchers.

- *Effect of PTC on Dispatcher* – The introduction of Positive Train Control will undoubtedly alter the dispatcher's job, just as it is expected to affect the locomotive engineer's job. With PTC the dispatcher's job could fundamentally change from one of actively dispatching trains and other track users to one of passively monitoring or supervising a closed loop system. Of particular concern is how PTC will affect a dispatcher's situation awareness, especially when time and safety-critical demands are made on the dispatcher. The impact of PTC on dispatcher workload, stress and fatigue should be examined in the early stages of PTC development so that the resultant PTC system can most effectively take into account the role of the dispatcher as system supervisor.

7. REFERENCES

Aschoff, J. (1979). Circadian Rhythms: General Features and Endocrinological Aspects. D. T. Krieger (ed.), *Endocrine Rhythms* (pp. 1-61). New York: Raven Press.

Barnes-Farrell, J. and Peotrowski, M. (1991). Discrepancies Between Chronological Age and Pesona Age as a Reflection of Unrelieved Worker Stress. *Work and Stress,* 5(3), 177-187.

Bartter, F. C., Delea, C. S., and Halberg, F. (1962). A Map of Blood and Urinary Changes Related to Circadian Variations in Adrenal Cortical Function in Normal Subjects. *Annals of the New York Academy of Sciences,* 98, 969-983.

Beck, A. (1987). *Beck Depression Inventory Manual.* Palo Alto, CA: Consulting Psychologists Press.

Bierbaum, C. R., Fulford, L. A., and Hamilton, D. B. (1990). *Task Analysis/Workload (TAWL) User's Guide* (Version 3.0 ed.). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Cooper, G. E., and Harper, R. P. (1969). *The Use of Pilot Ratings in the Evaluation of Aircraft Handling Qualities.* (Report No. NASA TN-D-5153). Moffett Field, CA: NASA Ames Research Center:

Czeisler, C. A., Kronauer, R. E., Allan, J. S., Duffy, J. F., Jewett, M. E., Brown, E. N., and Ronda, J. M. (1989). Bright Light Induction of Strong (Type 0) Resetting of the Human Circadian Pacemaker. *Science*, 244, 1328-1333.

Dahlgren, K. (1981). Long-term Adjustment of Circadian Rhythms to a Rotating Shiftwork Schedule. *Scandinavian Journal of Work, Environment, and Health,* 7, 141-151.

Dell'Erba, G., Venturi, P., Rizzo, F., Porcu, S., and Pancheri, P. (1994). Burnout and Health Status in Italian Air Traffic Controllers. *Aviation, Space, and Environmental Medicine,* 65, 315-322.

Devoe, D. B. (1974). *An Analysis of the Job of Railroad Train Dispatcher.* (Report No. FRA-ORD&D-74-37). Springfield, VA: National Technical Information Service.

Federal Railroad Administration. (1990). *National Train Dispatcher Safety Assessment.* Washington, D.C.: U.S. Department of Transportation.

Federal Railroad Administration. (1995). *National Train Dispatcher's Follow-up Review*. Washington, D.C.: U.S. Department of Transportation.

Fibiger, W., Evans, O., and Singer, G. (1986). Hormonal Responses to a Graded Mental Workload. *European Journal of Applied Physiology,* 55, 339-343.

Gordon, G., Tepas, D., Stock, C., and Walsh, J. (1979). Gaining Access to Shiftworkers Through Labor Unions. *Behavior Research Methods and Instrumentation,* 11, 14-17.

Gross, A. L., and Su, W. (1975). Defining a "Fair" or "Unbiased" Selection Model: A Question of Utilities. *Journal of Applied Psychology,* 60(3), 345-351.

Hart, S. G., and Staveland, L. E. (1988). Development of the NASA-TLX (Task Load Index): Results of Experimental and Theoretical Research. P. A. Hancock, & N. Meshkati (eds.), *Human Mental Workload* (pp. 139-183). Amsterdam: North Holland.

Hellhammer, D. H., Kirschbaum, C., and Belkien, L. (1987). Measurement of Salivary Cortisol Under Psychological Stimulation. J. N. Hintigen, D. H. Hellhammer, and G. Huppmann (eds.), *Advanced Methods in Psychobiology* (pp. 281-289). Toronto: Hogrefe.

Hunter, J. E., and Hunter, R. F. (1984). Validity and Utility of Alternative Predictors of Job Performance. *Psychological Bulletin,* 96(1), 72-98.

Johnson, L. C., and Naitoh, P. (1974). *The Operational Consequences of Sleep Deprivation and Sleep Deficit.* (Report No. AGARD-AG-193, NATO). London: Technical Editing and Reproduction.

Kirschbaum, C., Read, G. F., and Hellhammer, D. H. (1992). *Assessment of Hormones and Drugs in Biobehavioral Research.* Seattle: Hogrefe & Huber.

Kirschbaum, C., Read, G. F., and Hellhammer, D. H. (1994). Salivary Cortisol in Psychoneuroendocrine Research: Recent Developments and Applications. *Psychoneuroendocrinology,* 19(4), 313-333.

Knauth, P., and Rutenfranz, J. (1987). Shiftwork. J. M. Harrington (ed.), *Recent Advances in Occupational Health* (pp. 263-281). Edinburgh: Churchill Livingstone.

Laudat, M. H., Cerdas, S., Fournier, C., Guiban, D., Guilhaume, B., and Luton, J. P. (1988). Salivary Cortisol Measurement: A Practical Approach to Assess Pituitary-Adrenal Function. *Journal of Clinical Endocrinology and Metabolism,* 66(2), 343-348.

Matthews, G., Jones, D. M., and Chamberlain, A. G. (1990). Refining the Measurement of Mood: The UWIST Mood Adjective Checklist. *British Journal of Psychology,* 81, 17-42.

McConnell, E. (1998). Automated Vital Signs Monitoring Devices. *Nursing Management,* 29(2), 49-517.

McCord, C. P. (1948). Life and Death by the Minute. *Industrial Medicine,* (October), 377-382.

Menotti, A., and Seccareccia, F. (1985). Physical Activity at Work and Job Responsibility as Risk Factors for Fatal Coronary Heart Disease and Other Causes of Death. *Journal of Epidemiology and Community Health,* 39, 325-329.

Migeon, C. J., Tyler, F. H., Mahoney, J. P., Florentin, A. A., Castle, H., Bliss, E. L., and Samuels, L. T. (1956). The Diurnal Variation of Plasma Levels and Urinary Secretion of 17-hydroxycorticosteroids in Normal Subjects, Night Workers, and Blind Subjects. *Journal of Clinical Endocrinology and Metabolism,* 16, 622-633.

Moore-Ede, M. C., and Richardson, G. S. (1985). Medical Implications of Shiftwork. *Annual Review of Medicine,* 36, 607-617.

Nadler, E., Haines, M., Bonin, J., and Grossberg, M. (1997). Laptop Recording of Operational Taskload. *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

Nichols, C. T., and Tyler, F. H. (1967). Diurnal Variation in Adrenal Cortical Function. *Annual Review of Medicine,* 18, 313-324.

Peterson, R. E. (1957). Plasma Corticosterone and Hydroxycortisone Levels in Man. *Journal of Clinical Endocrinology and Metabolism,* 17, 1150-1157.

Pincus, G. (1943). A Diurnal Rhythm in the Excretion of Urinary Ketosteroids by Young Men. *Journal of Clinical Endocrinology and Metabolism,* 3, 195-199.

Reid, G. B., and Nygren, T. E. (1988). The Subjective Workload Assessment Technique: A Scaling Procedure for Measuring Mental Workload. P. A. Hancock, & N. Meshkati (eds.), *Human Mental Workload* (pp. 185-213). Amsterdam: North Holland.

Riad-Fahmy, D., Read, G. F., Walker, R. F., and Griffith, K. (1982). Steroids in Saliva for Assessing Endocrine Function. *Endocrinology Review,* 3, 367-395.

Roscoe, A. H. (1987). *In-flight Assessment of Workload Using Pilot Ratings and Heart Rate*. (Report No. AGARD-AG-282). NATO.

Sauter, S. and Murphy, L. (1995). *Organizational Risk Factors for Job Stress*. Washington, DC: American Psychological Association.

Schmidt-Reinwald, A., Pruessner, J. C., Hellhammer, D. H., Federenko, I., Rohleder, N., Schurmeyer, T. H., and Kirschbaum, C. The Cortisol Response to Awakening in Relation to Different Challenge Tests and a 12-Hour Cortisol Rhythm. *Life Sciences,* 64(18), 1653-1660.

Selye, H. (1978). *The Stress of Life* (Revised ed.,). New York: McGraw-Hill.

Spielberger, C. D. (1983). *State Trait Anxiety Inventory, Form Y*. Palo Alto, CA: Consulting Psychologists Press.

Tepas, D., and Carlvalhais, A. (1990). Sleep Patterns of Shiftworkers. (pp. 199-208). Philadelphia, PA: Hanley & Belfus, Inc.

Tepas, D., Walsh, J., Moss, P., and Armstrong, D. (1981). Polysomnographic Correlates of Shift Worker Perofmrance in the Laboratory. A. Reinberg, & N. A. P. Vieux (Editors), *Night and Shift Work: Biological and Social Aspects.* Oxford: Pergamon Press.

Tepas, D., and Monk, T. H. (1987). Work Schedules. G. Salvendy (ed.), *Handbook of Human Factors and Ergonomics* (Second ed.,). New York: John Wiley and Sons.

Tepas, D., Paley, M. J., and Popkin, S. M. (1997). Work Schedules and Sustained Performance. G. Salvendy (ed.), *Handbook of Human Factors and Ergonomics* (Second ed.). New York: John Wiley and Sons.

Thomas, G. R., Raslear, T. G., and Kuehn, G. I. (1997). *The Effects of Work Schedule on Train-Handling Performance and Sleep of Locomotive Engineers: A Simulator Study.* Washington, D.C.: FRA.

Vanderhorst, J. (1990). *A Comparison of Voice and Data Link Communication: Railroad Dispatcher's Perspective*. (Report No. PB91-130021INZ). Springfield, VA: National Technical Information Service.

Vining, R. F., and McGiley, R. A. (1986). Hormones in Saliva. *Critical Review in the Clinical Laboratory Sciences,* 23, 95-146.

Weiten, W. (1992). *Psychology Themes and Variations* (Second ed.). Pacific Grove, CA: Brooks/Cole.

Weitzman, E. D., Fukushima, D. K., Nogeire, C., Roffwarg, H., Gallagher, T. F., and Hellman, L. (1971). Twenty-four Hour Pattern of the Episodic Secretion of Cortisol in Normal Subjects. *Journal of Clinical Endocrinology and Metabolism,* 33, 14-22.

Wever, R. (1979). *The Circadian System of Man. Results of Experiments Under Temporal Isolation.* New York: Springer-Verlag.

Zeier, H. (1994). Workload and Psychophysiological Stress Reactions in Air Traffic Controllers. *Ergonomics,* 37(3), 525-539.

Zuckerman, M., and Lubin, B. (1986). *Multiple Affect Adjective Checklist - Revised.* San Diego: Educational and Industrial Testing Service.

APPENDIX A

DEFINITION OF ACTIWATCH DATA

Sleep Start – Time of sleep onset.

Sleep End – Time of sleep termination.

Bed Time - Time at which subject went to bed or turned off the lights.

Get up Time - Time at which subject left the bed or turned on the lights.

Time in bed – Difference in hours and minutes between the Bed Time and Get up time.

Estimated sleep time – Difference between the Sleep End and the Sleep Start times.

Sleep latency – Period of time required for sleep onset after going to bed. It is computed as the period between Bed Time and Sleep Start.

Sleep efficiency – An index of the amount of time in bed that is actually spent sleeping. Compute by dividing the Actual Sleep Time by the Time in Bed and multiplying by 100.

Number of sleep bouts – A count of continuous blocks of sleep.

Mean length of sleep bouts – The average length of the blocks of continuous sleep. Calculated by dividing the Actual Sleep Time by the Number of Sleep Bouts.

Movement and fragmentation index – An index of restlessness. Calculated by summing the Number of Minutes Moving Percentage with the Immobility Phases of 1 Minutes Percentage.

Number of Minutes Moving Percentage – Calculated by dividing the number of Minutes Moving by the Estimated Sleep Time and multiplying by 100.

Immobility Phases of 1 Minute – Number of one-minute increments in which there is no motion.

APPENDIX B

FORMS FOR FIELD STUDY

NOTE:
This appendix is not available in pdf format. The printed version of this report contains this appendix.

APPENDIX C

PILOT STUDY

The purpose of the pilot study was to test and evaluate the utility and practicality of a number of data collection measures and procedures for use in a railroad dispatching office. The pilot study was also used to identify potentially problematic situations that can occur in a field study of this nature, and means of addressing those situations. Data from the pilot study were used to evaluate the success of the data collection instruments and their protocols. These data were not used to draw conclusions about railroad dispatcher workload, stress and fatigue.

C.1 Materials

A number of test instruments were used in the pilot study. They included:

1. A background survey that contained questions pertaining to railroad dispatcher health and well-being; work scheduling and sleep habits; sources of stress in the workplace; interpersonal relationships; control over work; and quality of life.

2. A participant instruction and survey booklet that included an introduction to the study; definitions of workload, stress and fatigue; instructions for completing the daily subjective ratings; a set of five subjective workload, stress, and fatigue ratings for each day of the study; one sleep log for every day; and a debriefing survey.

3. Wrist-worn actigraphs.

4. Salivary cortisol assays.

5. A laptop computer-based data collection system.

6. An observation-based workload assessment system.

In addition, an informed consent form was developed for administration in the pilot study. The informed consent form described in detail the conditions of the study, as well as participants' rights and responsibilities.

C.2 Pilot Study Sites and Procedures

Three dispatching centers were selected as pilot study sites. These centers were located in three different geographical regions across the country and represented both types of railroad operations—freight and commute/passenger. All three pilot studies were conducted in the Fall of 1997. Table C-1 provides some details on the pilot study locations, dates, and procedures tested.

Participants read the instructions and signed an informed consent form prior to the start of data collection. It was emphasized that participants' primary functions of dispatching and safety were to take precedence over any study activities, and that if any conflict should arise, participants should complete their primary job duties first, and address the study tasks at a convenient time later.

C.2.1 Pilot Study No. 1

The first of the three pilot studies was conducted in a freight railroad dispatching center located in the Mid-Atlantic. The dispatching center was dimly lit, and each dispatcher worked at his/her own workstation (see Figure C-1). The instruments tested at this location included the background survey, daily subjective ratings, and the laptop computer-based modified NIMS workload data collector. Two participants were recruited and were paid $35 each for their participation. The primary goal for this pilot study was to explore the feasibility of using the modified NIMS data collector to collect railroad dispatching-related information *in situ*. Data were collected for two days, following two days of experimenter practice with the laptop PC data collection system.

Problems using the NIMS data collector were apparent at the outset of the pilot study. These problems are described in subsection C.3.1. As a consequence, the possibility of collecting frequencies of various activities (e.g., number of incoming versus outgoing calls, number of Form D's completed, etc.) by hand was explored as an alternative to the computer-based system. Frequency counts are not sensitive to time since they can be recorded after their occurrence using computer- and paper-based records, and they provide similar information to what the PC system could yield. Due to the difficulties in using the computer-based system to collect dispatcher-related task activity at the freight operation, plans for the second pilot study were modified so that frequency count information would be incorporated into the design of the pilot study.

Table C-1. Pilot study locations, dates and procedures

Study No.	Location	Type of Operation	Procedures Tested
1.	Mid-Atlantic	Freight	Background survey, subjective ratings, sleep logs, NIMS
2.	New England	Commuter/passenger	Actigraph, background survey, salivary cortisol, sleep logs, subjective ratings, objective workload measures
3.	Midwest	Commuter	mTAWL

Figure C-1. Dispatching desk at pilot site No. 1

C.2.2 Pilot Study No. 2

The second pilot study was conducted at a passenger railroad operation located in New England. The dispatching center was also dimly lit, and situated in an amphitheater-style setting. All of the test instruments except for the mTAWL were explored in this pilot study. Five participants participated for seven consecutive days. Railroad management at the dispatching office assisted in soliciting participants. Participants were each paid $100 for their participation. Only first shift dispatchers, those working 7 am to 3 pm, were recruited to simplify data collection procedures.

Background surveys, informed consent statements, survey booklets, and actigraphs were all distributed shortly before the start of the data collection period. To protect the identity of the participants, each was assigned an unique identification number. Participants were then provided with explicit instructions for completing each of the test instruments. The background survey was to be completed at home and returned to the experimenter as soon as was convenient. Participants were asked to wear the actigraphs for the entire week except when showering/bathing/swimming, and were instructed to complete a sleep log just after rising each morning. Subjective rating forms were to be completed every 2 hr, starting at the beginning of the participant's shift (baseline). In addition to the subjective rating forms, participants were asked to record the number of trains and the number of additional (i.e., non-train) track users that they had handled over the last 2 hr.

As part of the procedure for collecting participants' saliva, participants were asked to refrain from eating, drinking and smoking for 20 min prior to each saliva sampling. At the time of sampling, the dispatcher placed a small, sterile cotton cylinder, similar to those used by dentists, in their mouth without touching the material. Once soaked with saliva, about two minutes, the participant placed the cotton into a sterile plastic tube, sealed it, and placed it in an ice chest located at their desk as part of the study. Participants were asked to write down the time that they completed taking the sample as well as the unique number of the sample. Participants provided a saliva sample five times daily, upon starting their shift and every 2 hr thereafter. At the end of each day, technicians collected the saliva samples and delivered them to a local lab for analysis.

An experimenter also observed each participant once for an entire shift to collect frequency count information (e.g., number of incoming versus outgoing calls, etc.). At the end of the study period, each participant completed the debriefing survey, and returned the background surveys, the survey booklets containing the daily subjective ratings, sleep logs and debriefing survey, and the actigraphs. Participants were then thanked and paid for their participation. Data from dispatcher records pertaining to the number of Form Ds and foul time permits issued were collected after the pilot study was completed.

While most data were analyzed in-house, the Actigraph data were downloaded and analyzed by Ambulatory Monitoring personnel.

C.2.3 Pilot Study No. 3

The third pilot study was conducted at a commuter railroad operation located in the Midwest. The dispatching center was located in a small, modern facility. The goal of this pilot study was to evaluate the modified TAWL (mTAWL) approach to assessing a railroad dispatcher's workload.

In consultation with site personnel, one dispatching desk was selected for observation during evening rush hour. This desk was considered appropriate for observation since it had a very high level of traffic and involved both scheduled commuter and unscheduled freight trains. The first observations were made with a floor supervisor present to answer questions and assist in the creation of task definitions and resource channel loads for the mTAWL tasks. Three 2 to 3 hr observations were conducted to determine task definitions and channel loadings.

Preliminary observations indicated that the channel loads associated with the performance of any given task did not fluctuate over time in any predictable way. Consequently, average channel loadings were computed for each task. Table C-2 presents the subcategories and average channel loadings for each type of task.

During a single 2-1/2 hr observation period, the observed dispatcher handled 118 track occupants, with a majority of these occupants present during the 2 hr evening rush home. This level of activity makes it impossible to simultaneously observe, record, and classify the dispatcher's behaviors. With multiple events occurring in a 1 min period, it was deemed

Table C-2. Task types and channel loadings

Task Area	Task Description	Channel Loading			
		Audio	Visual	Cognitive	Psycho-motor
Background	• *Auditory monitoring* - listening for radio or telephone calls.	1	0	0	0
	• *Visual monitoring* - watching for unanticipated events on the informational displays.	0	1	0	0
	• *Background "phone"* - telephone or radio calls made during the course of work at times chosen by the dispatcher and not directly related to the anticipated or actual control of a specific train or work crew.	3	1	2	1
	• *Background "computer"* - computer or paper entries required during the shift that do not require an immediate action and can be performed at times chosen by the dispatcher.	0	2	2	6
Foreground	• *Foreground "phone"* – telephone or radio calls made or received during the course of work which required an immediate response and which were not directly related to previously anticipated actions or the current control of a specific train or work crew.	3	1	2	1
	• *Foreground "computer"* – computer or paper entries required during the shift that were not directly related to previously anticipated actions or the current control of a specific train or work crew but that had to be taken care of immediately.	0	2	2	6
Anticipatory	• *Anticipated unopposed* – the anticipated arrival of a track occupant whose progress toward the dispatcher's territory is expected to be unimpeded.	1	1	1	0
	• *Anticipated opposed* – the anticipated arrival of a track occupant whose progress to the dispatcher's territory was expected to be opposed by another track occupant or occupants.	1	1	2	0
Track Management	• *Track occupant unopposed* - a period of time during which a specific track occupant under the dispatcher's control is unimpeded.	1	1	1	1

appropriate to observe and record the events as they occurred, and later classify them as a post-processing activity. Train sheets were especially useful in the categorization process to verify track occupancy times and locations after-the-fact.

Observed events were then entered into a time-sequenced spreadsheet that also included a graphical representation of train movements and track occupancy. Columns in the spreadsheet were used to represent different railroad tracks, and were arranged to represent spatial relationships, so that conflicts in train movement and track occupancy would be readily apparent. This spreadsheet also revealed the dispatcher's responses to triggering events and how track occupancy conflicts were resolved. A sample spreadsheet is provided in Figure C-2.

C.3 Pilot Study Results

This section describes the findings and experiences from the three pilot studies. The purpose of the pilot studies was to evaluate the feasibility and appropriateness of using the selected candidate data collection instruments and procedures in the railroad dispatching work environment. Consequently, results are reported in terms of how successful each of the test instruments or procedures was with respect to field setting and the nature of the dispatcher's job. Results from the pilot studies were then used to modify the testing instruments and procedures used in the full field study.

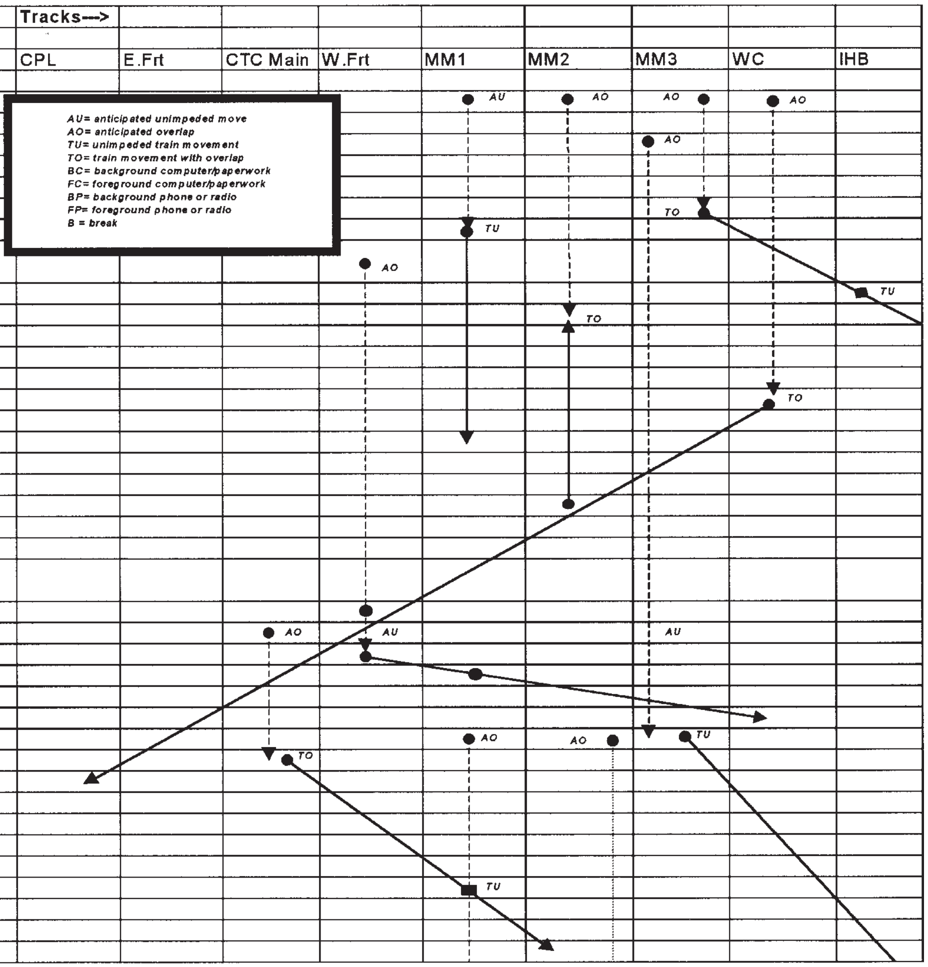
C.3.1 Results from Pilot Study No. 1

The NIMS data collector proved very difficult and inefficient to use for several reasons. First, the dynamic nature of the dispatcher's work made it difficult to determine when activities began and ended. For example, calls are put on hold/standby and dispatchers switch back and forth between activities (e.g., from recording something in their train sheet, to changing a switch/signal on their CTC, to returning to the train sheet to complete their entry, etc.). Second, often it took several seconds to determine what activity was taking place, thereby losing valuable recording time. For example, it took time to determine if a call or radio communication was incoming or outgoing. Furthermore, some conversations were very brief. As a result, it was easy to lose recording time as well as to miss entire events. Third, many dispatcher activities were also of short duration, high frequency and highly repetitive, making it very difficult to accurately capture all of the activities taking place. Finally, individuals, including the train dispatcher under observation, attempted to engage the observer in conversation, making it difficult to make accurate (i.e., timely) recordings.

For all of these reasons, the NIMS data collection system was abandoned. Instead, simple task frequency information (e.g., number of incoming versus outgoing calls, number of Form D's completed, etc.) was collected. It was both feasible and unobtrusive to collect this information. No attempt was made to analyze either the NIMS data or the task frequency data since data was collected only on two 8 hr shifts.

Separately, feedback from the two participants indicated no difficulties or problems with completing the background survey, daily subjective ratings, or sleep logs.

11-Dec-98				
TIME	TRAIN	TRACK	DIR	EVENT
1559				Phone; gets mark up WC T049
1559				Enters trains into computer stack (FC)
1601				Radio; sets up CP1399
1603				
1605				Radio; calls CP So Side to advise move
1606	*CP So Side*	***3 to NPB***	*West*	***Enters circuit***
1607	***MW 2223***	***1 to 1***	*West*	***Enters circuit***
1608				Phone; calls dispatcher to set up time for WCCXFFLD at 1655
1608				Radio; calls MW2223 for location; no answer
1609	***MW2242***	***2 to 2***	*East*	***Enters circuit***
1610	*CP So Side*	***3 to NPB***	*West*	***Leaves circuit***
1613				Makes entry to late reports (BC)
1614				OS report (BC)
1616	***WCT 049***	*WC to CPL*	*West*	***Enters circuit***
1616				Phone; line up
1617	***MW 2223***	***1 to 1***	*West*	***Leaves circuit***
1617				OS report (BC)
1617				Hand sets part of a route
1619	***MW2242***	***2 to 2***	*East*	***Leaves circuit***
1619				OS report (BC)
1619				Hand sets part of a route
1620				Computer drops stacked routes for unknown reason; resets. (FC)
1620				Radio: lineup for WCCXFFLD; revised.
1621				Radio to NC112 for location
1623	*WCCXFFLD*	*WFL to WC*	*East*	***Enters circuit***
1623				Radio to WC T049; on move? (switching)
1626				Radio to CP1399; tells them to watch for signal
1628	*WCCXFFLD*	*WFL to WC*	*East*	***Leaves circuit***
1628	***CP1399***	*3 to NPB*	*West*	***Enters circuit***
1629	***NC112***	*CTC to 2*	*East*	***Enters circuit***
1630	***WCT 049***	*WC to CPL*	*West*	***Leaves circuit***
1631				Attempts to phone in OS times; busy (BR)
1632				Phone; line up
1634				Radio; calls for time on MW2244 (late)
1634				SWS11 is on approach (au)
1635				MW2244 on approach
1635				Reports MW2244 late
1637				Partially lines route for SWS11 (short window)
1637	***NC112***	*CTC to 2*	*East*	***Leaves circuit***



436-FRA-97120-7

Figure C-2. Sample mTAWL spreadsheet

C.3.2 Results from Pilot Study No. 2

Results from the second pilot study are organized around the particular instrument that was being tested and evaluated in the pilot study.

C.3.2.1 Background Survey

Based on casual conversations with participants, surveys did not take long to complete, nor were the contents difficult to understand. Data from the background surveys were examined to determine the quality and range of responses to the various question sections. Among the more interesting responses are the following: 1) four out of five dispatchers had experienced or were currently experiencing health symptoms related to stress, for example, back and neck pain, intestinal upsets, anxiety, and headaches. 2) In terms of work scheduling, all of the dispatchers complained that they were too busy to take breaks and that it was difficult to get another dispatcher to cover for them, even in emergencies. In addition, most complained about the lack of a regular lunch break and the attendant necessity of having to eat hurriedly at their workstations. 3) Four out of the five dispatchers reported at least one symptom of insomnia and experienced daytime sleepiness as a consequence. 4) Most of the participants rated the visual demands of their work as high or very high, and also reported symptoms of eyestrain. 5) Lack of control over features of their work and working environment were reported as important sources of occupational stress. 6) The occasional loss of temper on the job was nearly universal among surveyed volunteers, and many also confided that they tend to take their work stress home with them at the end of the day. The items on the background survey thus seemed sufficiently sensitive to be used in the full field study.

C.3.2.2 Objective Workload

Objective workload data was collected in the form of number of trains and other track users that the participants handled, the number of mandatory directives (i.e., track usage authorities) issued and/or canceled by dispatchers, and the number of incoming and outgoing phone calls. The data were collapsed across participants as well as days, and are illustrated in Figures C-3 through C-6. Figure C-3 shows that train traffic peaked early in the shift, then decreased after morning rush hour.

There was a corresponding increase in other track usage during mid-morning, as maintenance crews made track repairs and inspections during periods of light train traffic. This pattern is also reflected in increases in mandatory directives (Figure C-4) and phone calls (Figure C-5) following rush hour.

Inspection of the data indicates that dispatcher-reported data and official dispatcher records are a viable source of workload information. However, there were several potential problems that had implications for the field study. First, dispatchers at this pilot study site were not required to report when foul time permits were issued; only when they were voided. Furthermore, it would not be feasible to monitor all the participants in the full field study in order to collect counts on the number of phone and radio calls and the number of Form D's issued and

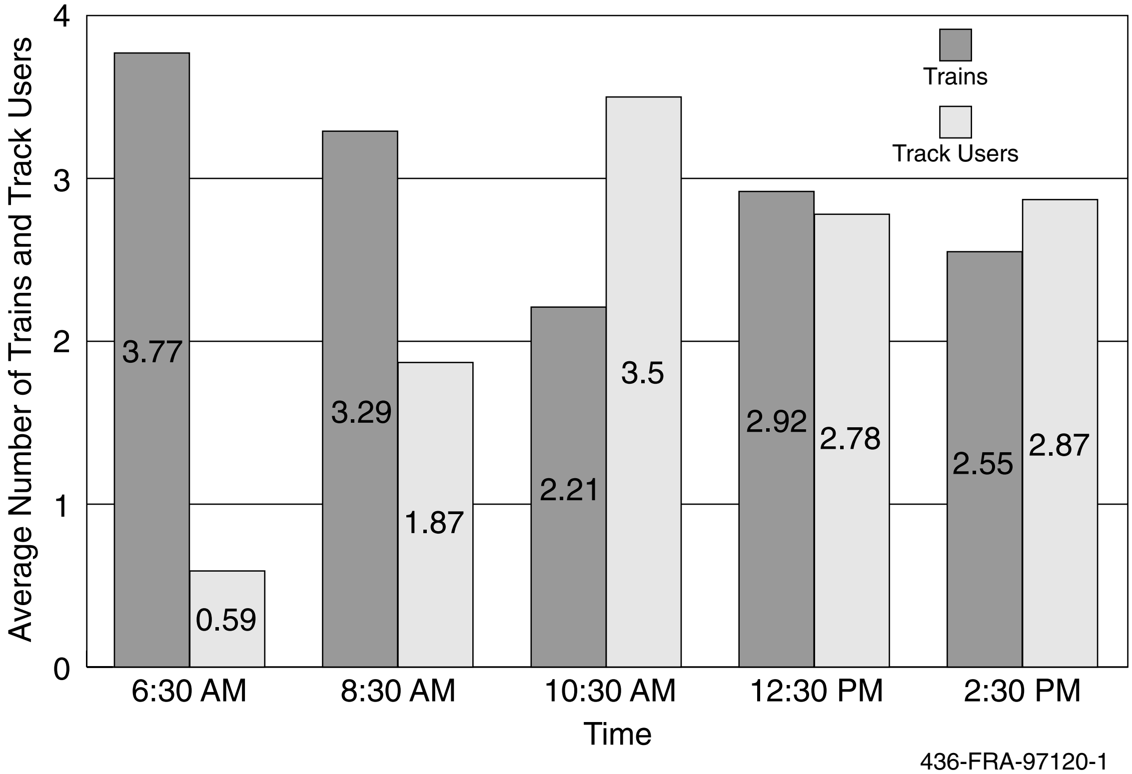


Figure C-3. Traffic volume

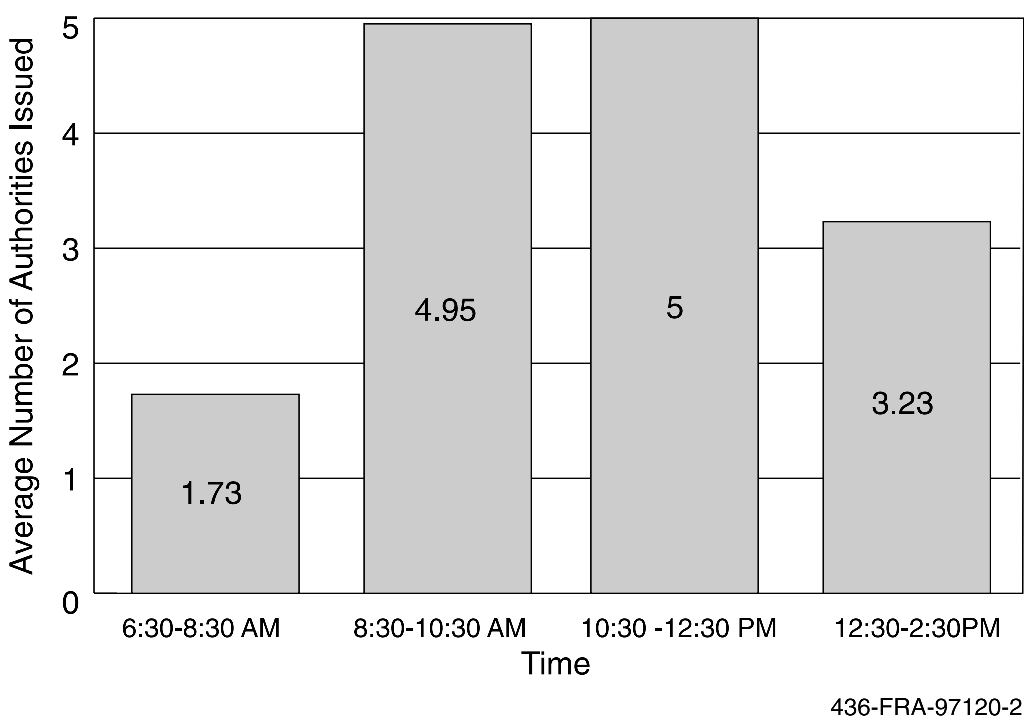


Figure C-4. Track usage authorities

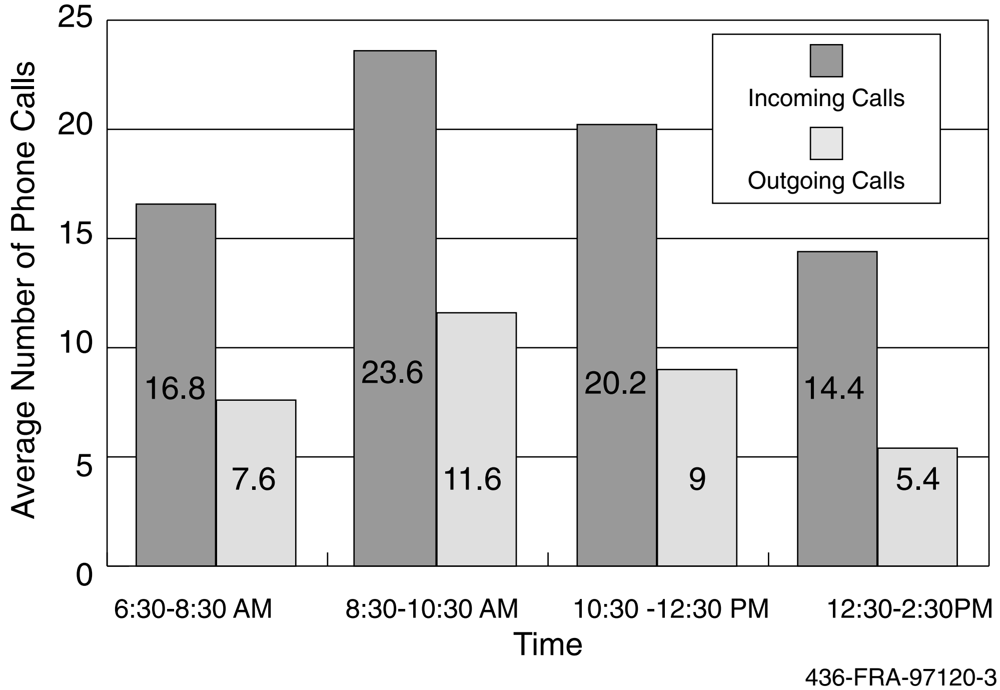


Figure C-5. Phone calls

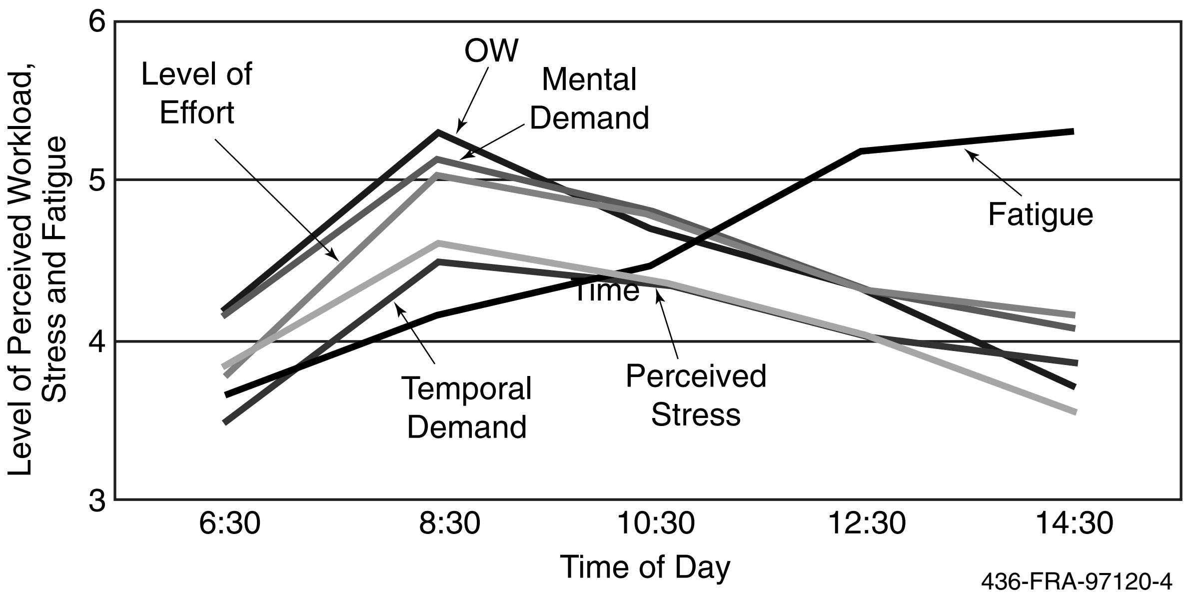


Figure C-6. Subjective ratings

cancelled. Given this, a practical alternative was devised for the field study whereby Form D issuance and cancellation information would be collected at a future time once this paperwork had been processed. A technical solution to counting the number of telephone and radio calls was also sought, but none proved viable. Consequently, information on the number of incoming and outgoing phone and radio calls would not be collected in the full field study.

C.3.2.3 Subjective Ratings

Figure C-6 shows that, collapsed across all five volunteers, fatigue gradually increased throughout the day while workload and stress rose to a peak at the height of morning rush hour and then steadily declined. The three workload subscales followed the same pattern as overall workload. Each volunteer's overall workload ratings were also correlated with the number of track users (trains and others) on his territory at five intervals during the day for each day of study participation. Pearson product-moment correlations were positive and strong, ranging from 0.57 to 0.83. The strong association with the objective workload measure suggests that the dispatchers' self-ratings are an accurate reflection of task demands, an important issue in establishing the validity of self-report instruments.

Participants completed all of the subjective rating forms, suggesting that the forms did not take too much of the dispatcher's time to complete. Participants had several critical comments, however: First, several participants noted that the survey booklets were large and unwieldy. Second, participants noted that these surveys could actually increase workload when a dispatcher is busy; third, it was felt that it was sometimes difficult to evaluate one's own stress level on the spot. And finally, one participant felt that recording the number and types of trains and other track users on their territories every 2 hr was time-consuming.

Based on this feedback, the survey booklets that contained all of the sleep logs and subjective ratings for the entire week were replaced with daily rating sheet packets. These packets were dated, and could be picked up and dropped off at the dispatching center. This change would provide two benefits. First, since the new rating packet consisted of only a few pages, it would not be a problem to keep it on his/her desk, and participants would not have to keep track of it over the course of the study. Second, participants would not be able to review the preceding days' ratings, a situation that could influence current subjective ratings.

C.3.2.4 Cortisol Measurement

The saliva sampling procedures, as well as the assays, were highly successful. Over 95 percent of these samples were collected and yielded useful data, indicating that this measure was appropriate and feasible for a workplace study of stress.

Participants did have some concerns and suggestions concerning this measure, however. In terms of the saliva collection procedure itself, participants expressed a reluctance to restrict their food intake and smoking prior to providing a saliva sample. Since this is necessary in order to get an accurate reading, information regarding these restrictions was explicitly supplied during the recruitment of participants for the field study. It was felt that this information would better prepare the participant to follow the procedure, and would provide him/her with the opportunity to decline to participate if this restriction were too inconvenient. Separately, one participant commented that chewing on the cotton swab interfered slightly with his communications, and another participant felt that the pressure to produce the sample could increase workload. There was also a concern that the timing of the saliva collection was sometimes inconsistent with high workload periods. Since most participants did not find the salivary cortisol procedure too invasive, however, and because it was the only physiological measure of stress incorporated into the study, salivary cortisol would also be used in the full field study.

C.3.2.5 Actigraphy and Sleep Logs

Inspection of participants' sleep logs indicated that four of the five were sleeping an average of 6 hr per night, including days off. Furthermore, it was not uncommon for some dispatchers to occasionally get only 4 to 5 hr of sleep per night. Subjective reports of sleep duration were generally consistent with the results of the actigraphy (a mean of 6.25 hr per night). In a few cases, however, the actigraphs either overestimated or underestimated the participants' sleep duration. With respect to overestimation of sleep, the actigraph was apt to score a still but wakeful subject as being asleep. In terms of underestimating sleep length, a participant may be tossing and turning in Stage I (light) sleep, but the actigraph data will indicate that s/he was awake. This is a problem with using motion to infer sleep/wakefulness, not just with the particular actigraph chosen for the study.

One other problem that arose with the actigraphy related to the physical size of the actigraph. Many of the participants felt that it was rather large, heavy, and uncomfortable. This was a particular concern to the participants if the watch was to be worn for two consecutive weeks, as in the field study. Based on this feedback, alternative wrist-worn actigraphs were explored for the field study, and a new watch design was eventually selected that addressed the pilot study participants concerns.

There was one participant whose subjective reports and actigraph conclusions were dissociated. This dispatcher was an extremely restless sleeper, who regularly reported getting much more sleep than was indicated by his actigraph recordings. He also indicated that he used alcohol as a relaxant and sleep aid. This presents a number of problems. First, frequent awakenings to toilet are likely. Second, alcohol tends to diminish Stages III and IV (deep) sleep, making sleep less restful and restorative. Finally, alcohol has a tendency to affect perceptions, making the individual less accurate at estimating or recalling his previous night's sleep. In this pilot study, as in many cases of dissociation, it was difficult to tell whether the actigraph recordings or subjective reports were more accurate.

Regarding the sleep logs, again there was high compliance in completing all the items on these forms over the entire week. The data provided by these forms appeared to have face validity. All but one of the dispatchers reported waking up an average of 1 to 3 times per night, though some awoke as frequently as 5 to 6 times per night. Four out of five volunteers reported no use of sleep aids, while the fifth dispatcher admitted to using alcohol as a sleep aid. Ease of falling asleep and ease of getting up ranged generally from very easy to fairly difficult, with a few reports of "very difficult." Ratings of sleep sufficiency were split fairly evenly between sufficient and insufficient, with only one report each of wholly insufficient and more than sufficient. Assessments of overall sleep quality ranged from fair to good for a majority of the nights (74 percent of the nights, or 26 out of 35 rest periods). Reports of overall fatigue generally ranged from somewhat fatigued to fairly rested, with few reports of being either very fatigued or very rested.

C.3.2.6 Debriefing Survey

This survey served as a mechanism for the participants to provide feedback about the various methodologies used during the pilot study. In general, the tested methodologies and protocol did not compromise the participants' ability to perform their daily work activities and routine. All participants expressed a willingness to participate again in the longer field study, with appropriate financial remuneration. Participants felt that $100 per week was fair compensation for the amount of involvement required of study participants. The debriefing survey was also used to collect additional information such as the date of the most recent day(s) off immediately prior to the start of the field study.

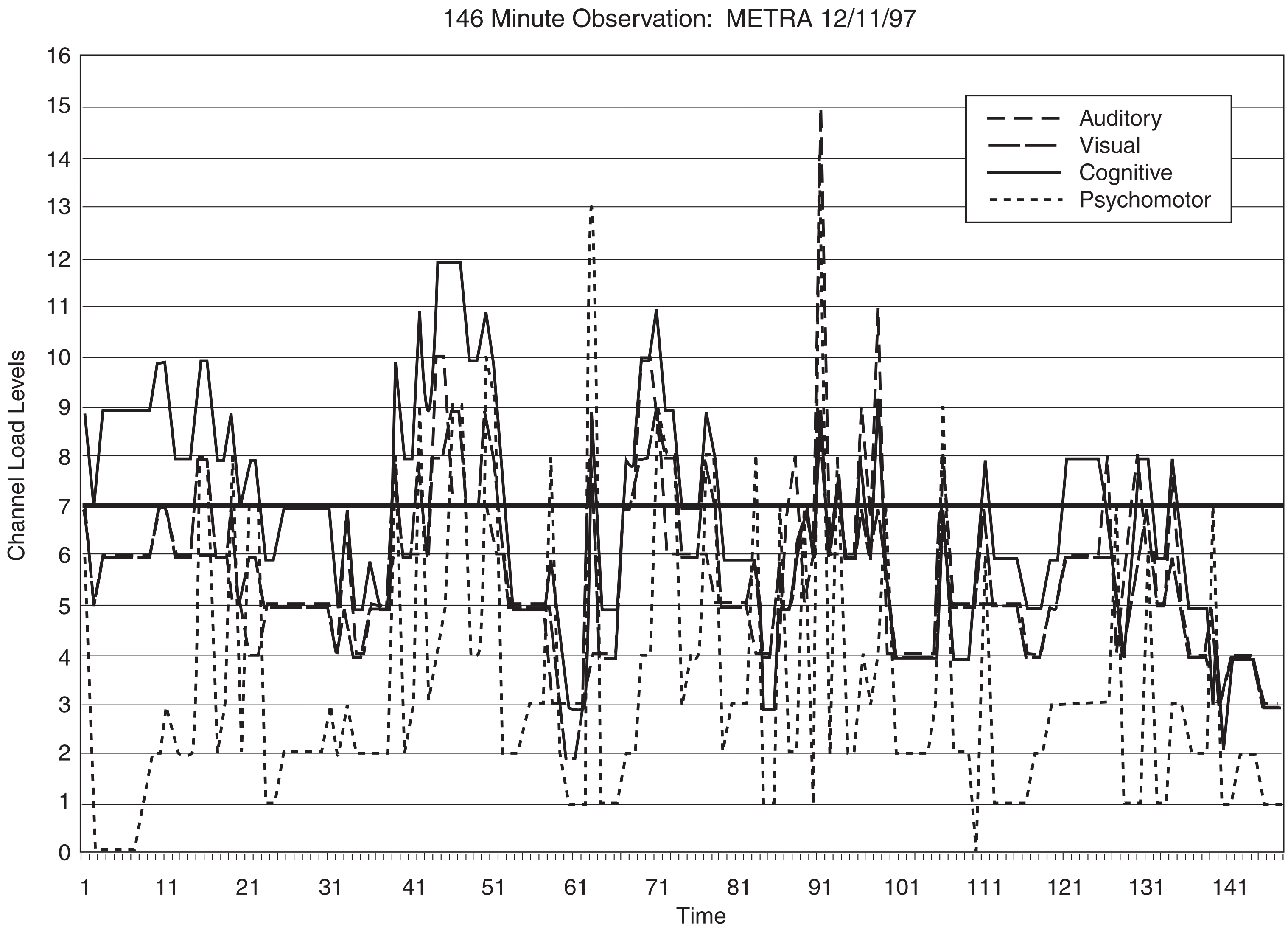
C.3.3 Results from Pilot Study No. 3

The adoption of the TAWL approach from a military flight deck to railroad dispatching center was successful. The task load totals for each channel were used to generate Figure C-7, where the four lines represent Auditory, Visual, Cognitive, and Psychomotor loads. The duration of the observation was 146 min. Time progresses from left to right on the abscissa. The height of the lines represents resource channel loads. When a line exceeds the value of 7 on the vertical scale, it implies an overload for that channel.

The TAWL concept of maximum channel load (7) implies that an operator's ability to perform a task may be compromised if a channel is loaded beyond that point. Figure C-7 shows a number of overload periods. The most significant of these occurs between the 41st and 51st min. During this period, all channels were overloaded. An investigation of the written record shows that six trains were present during this time period and considerable effort was expended in successfully moving them, particularly since many of the movements were potentially opposed.

During the first 20 min of the session, observed overload primarily involved the cognitive channel. The written record shows that a long, slow moving freight train unexpectedly appeared and had to be moved across the high-speed passenger main lines. The move was accomplished without delay or incident, but the overload of the cognitive channel implies a high level of planning and mental effort.

Based on this information, the mTAWL methodology was incorporated into the design of the full field study.



536-FRA-00045-1

Figure C-7. Channel loadings

APPENDIX D

mTAWL RESULTS

A total of 22 mTAWL observations were done, 10 at the freight operation and 12 at the passenger operation. The mTAWL data for six of the freight observations are presented in subsection 5.3.1. This appendix contains the data for the remaining four freight observations along with the data for the 12 passenger observations The last four graphs display the raw mTAWL data for two passenger and two freight observations using a 15 min moving average.

NOTE:
This appendix is not available in pdf format. The printed version of this report contains this appendix.

APPENDIX E

SUPPORTING STATISTICAL ANALYSES

Table E-1. Post hoc (Tukey HSD) comparisons by shift for each workload variable by study site

Site	Variable	Day versus Evening	Evening versus Night	Night versus Day
Freight	No. Trains 2h	**p=0.004**	**p=0.030**	Ns
	No. Users 2h	**p=0.010**	**p<0.001**	**p<0.001**
Passenger	No. Trains 2h	Ns	Ns	Ns
	No. Users 2h	Ns	Ns	Ns
Freight	No. Trains 4h	**p=0.004**	**p<0.001**	Ns
	No. Users 4h	**p<0.001**	ns	**p<0.001**
Passenger	No. Trains 4h	Ns	**p=0.009**	**p=0.001**
	No. Users 4h	**p<0.001**	ns	**p<0.001**
Freight	No. Trains 6h	Ns	**p=0.031**	Ns
	No. Users 6h	**p<0.001**	ns	**p<0.001**
Passenger	No. Trains 6h	Ns	**p=0.006**	**p=0.003**
	No. Users 6h	**p<0.001**	Ns	**p<0.001**
Freight	No. Trains 8h	Ns	Ns	Ns
	No. Users 8h	**p<0.001**	Ns	**p<0.001**
Passenger	No. Trains 8h	Ns	Ns	Ns
	No. Users 8h	**p=0.027**	Ns	**p=0.001**

Table E-2. Mean subjective ratings of the workload, stress and fatigue filled in at the beginning of shift and at 2 hr increments grouped by shift

Variable	Type of Operation	Day	Evening	Night	Overall
$Workload_b$	Freight	26.4	42.39	50.05	38.83
	Passenger	**40.40***	44.38	43.34	42.06
$Stress_b$	Freight	17.32	23.43	38.71	25.97
	Passenger	**25.42***	26.81	31.09	27.35
$Fatigue_b$	Freight	14.75	15.06	26.05	18.33
	Passenger	**24.16***	20.75	32.80	**25.32***
Workload 2h	Freight	35.88	43.62	43.15	40.54
	Passenger	**52.67***	55.27	33.20	**48.71***
Stress 2h	Freight	28.97	30.72	37.27	32.03
	Passenger	38.38	28.03	28.03	33.38
Fatigue 2h	Freight	18.57	17.71	33.27	22.68
	Passenger	**32.85***	20.97	38.31	**30.87***
Workload 4h	Freight	41.23	44.04	**43.95***	42.89
	Passenger	**53.52***	47.08	23.43	45.01
Stress 4h	Freight	36.90	31.47	**36.62***	34.97
	Passenger	40.20	29.60	23.2	33.77
Fatigue 4h	Freight	25.79	21.90	39.87	28.65
	Passenger	**37.78***	24.57	42.51	**35.56***
Workload 6h	Freight	40.76	37.33	**44.38***	40.57
	Passenger	46.69	47.95	32.91	43.59
Stress 6h	Freight	37.75	31.31	**38.93***	35.92
	Passenger	36.48	29.92	26.94	32.51
Fatigue 6h	Freight	30.43	26.37	44.38	33.26
	Passenger	**40.84***	28.70	48.16	**40.78***
Workload 8h	Freight	41.90	39.97	46.73	42.58
	Passenger	45.84	48.22	45.03	46.17
Stress 8h	Freight	42.14	32.72	41.71	38.71
	Passenger	37.57	34.75	40.25	37.40
Fatigue 8h	Freight	35.40	30.10	45.12	36.47
	Passenger	43.55	37.25	53.34	**44.01***
Overall Workload	Freight	37.19	41.15	**45.68***	
	Passenger	**47.81***	**48.58***	34.91	
Overall Stress	Freight	32.53	29.96	**38.64***	
	Passenger	35.66	29.79	29.34	
Overall Fatigue	Freight	24.91	22.26	37.64	
	Passenger	**36.27***	26.42	42.06	

*Indicates $p<0.01$ level of significance on an independent sample t-test.

Table E-3. Post hoc (Tukey HSD) comparisons by shift for each variable by study site

Type of Operation	Variable	Day versus Evening	Evening versus Night	Night versus Day
Freight	Workload 0h	**p=0.001**	Ns	**p<0.001**
	Stress 0h	Ns	**p<0.001**	**p<0.001**
	Fatigue 0h	Ns	**p<0.001**	**p<0.001**
Passenger	Workload 0h	Ns	Ns	Ns
	Stress 0h	Ns	Ns	Ns
	Fatigue 0h	Ns	Ns	Ns
Freight	Workload 2h	Ns	Ns	Ns
	Stress 2h	Ns	Ns	Ns
	Fatigue 2h	Ns	**p<0.001**	**p<0.001**
Passenger	Workload 2h	Ns	**p=0.002**	**p=0.003**
	Stress 2h	Ns	Ns	Ns
	Fatigue 2h	**p=0.034**	**p=0.010**	Ns
Freight	Workload 4h	Ns	Ns	Ns
	Stress 4h	Ns	Ns	Ns
	Fatigue 4h	Ns	**p<0.001**	**p<0.001**
Passenger	Workload 4h	Ns	**p<0.001**	**p<0.001**
	Stress 4h	Ns	Ns	**p=0.001**
	Fatigue 4h	**p=0.016**	**p=0.005**	Ns
Freight	Workload 6h	Ns	Ns	Ns
	Stress 6h	Ns	Ns	Ns
	Fatigue 6h	Ns	**p<0.001**	**p=0.002**
Passenger	Workload 6h	Ns	**p=0.038**	**p=0.033**
	Stress 6h	Ns	Ns	Ns
	Fatigue 6h	Ns	**p=0.003**	Ns
Freight	Workload 8h	Ns	Ns	Ns
	Stress 8h	Ns	Ns	Ns
	Fatigue 8h	Ns	**p=0.010**	Ns
Passenger	Workload 8h	Ns	Ns	Ns
	Stress 8h	Ns	Ns	Ns
	Fatigue 8h	**p=0.027**	**p=0.047**	Ns

Table E-4. Pearson correlations between subjective workload, stress and fatigue and reported workload by shift and hours into duty

Site	Shift	Number of Trains by Workload	Number of Trains by Stress	Number of Trains by Fatigue	Number of Track Users by Workload	Number of Track Users by Stress	Number of Track Users by Fatigue
Freight	Day	0.556*	0.407*	0.228	0.392*	0.364*	0.153
2 hr	Evening	0.554*	0.547*	0.201	0.254*	0.296*	0.188
	Night	0.606*	0.644*	0.299*	0.169	-0.019	-0.226
Passenger	Day	0.589*	0.281*	0.072	0.032	-0.111	-0.122
2 hr	Evening	0.642*	0.367*	0.272	-0.120	-0.009	-0.044
	Night	0.471*	0.322	0.120	0.471*	0.367*	0.101
Freight	Day	0.438*	0.255*	0.166	0.448*	0.379*	0.283*
4 hr	Evening	0.528*	0.502*	0.323*	0.395*	0.453*	0.318*
	Night	0.562*	0.585*	0.146	0.325*	0.483*	0.104
Passenger	Day	0.223*	0.039	0.098	0.202	0.008	-0.067
4 hr	Evening	0.374*	0.165	0.224	0.292	0.262	0.380*
	Night	0.409*	0.283	0.171	0.615*	0.604*	0.269
Freight	Day	0.463*	0.228	0.097	0.520*	0.396*	0.286*
6 hr	Evening	0.481*	0.530*	0.435*	0.111	0.095	-0.068
	Night	0.492*	0.353*	0.300*	0.344*	0.590*	0.004
Passenger	Day	0.185	0.057	0.057	0.017	-0.002	-0.121
6 hr	Evening	0.451*	0.269	0.268	0.558*	0.553*	0.476*
	Night	0.494*	0.428*	0.282	0.372*	0.291	0.097
Freight	Day	0.440*	0.336*	0.267*	0.474*	0.450*	0.395*
8 hr	Evening	0.534*	0.501*	0.358*	0.177	0.352*	0.143
	Night	0.648*	0.525*	0.096	0.115	0.108	0.014
Passenger	Day	0.361*	0.265*	0.160	-0.055	-0.062	-0.131
8 hr	Evening	0.320	0.225	0.384*	0.266	0.361*	0.291
	Night	0.653*	0.662*	0.347	-0.060	0.009	-0.253

*Indicates significance level of p<01.