



U.S. Department
of Transportation
**Federal Railroad
Administration**

THE EFFECTS OF WORK SCHEDULE ON TRAIN HANDLING PERFORMANCE AND SLEEP OF LOCOMOTIVE ENGINEERS: A SIMULATOR STUDY

Office of Research and Development
Washington, DC, 20590

Garold R. Thomas

Thomas G. Raslear

Federal Railroad Administration

George I. Kuehn

IIT Research Institute

DOT/FRA/ORD-97-09

**July 1997
Final Report**

This document is available
through the National
Technical Information Service,
Springfield, Virginia, 22161

cover

NOTICE

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

NOTICE

The United States Government does not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the objective of this report.

COVER

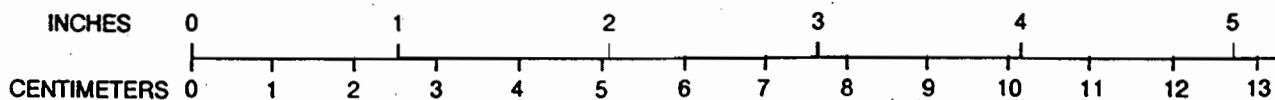
Technical Report Documentation Page

1. Report No.		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle The Effects of Work Schedule on Train-Handling Performance and Sleep of Locomotive Engineers: A Similar Study				5. Report Date 07/19/97	
				6. Performing Organizational Code	
				8. Performing Organization Report No.	
7. Author(s) G. R. Thomas (FRA), T. G. Raslear, G. I. Kuehn (IITRI)				10. Work Unit No. (TRAIS)	
9. Performing Organization Name and Address IIT Research Institute 10 West 35th Street Chicago, IL 60616-3799				11. Contract or Grant No. DTR53-82-C-00254, T.O. 10	
				13. Type of Report and Period Covered Final Report	
12. Sponsoring Agency Name and Address Federal Railroad Administration Office of Research and Development, Mail Stop 20 400 Seventh Street, SW Washington, DC 20590				14. Sponsoring Agency Code RDV-32	
				15. Supplementary Notes	
16. Abstract Current Federal regulations governing Hours of Service for locomotive engineers allow work schedules that have backwards rotating shift start times. Locomotive engineers who work under such schedules can accumulate a progressive sleep debt over a period of days. The present study demonstrates that schedules which have these characteristics are easily composed and do, indeed, result in sleep durations which are considerably less than those obtained by the general population. The locomotive engineers in this study, while working on such schedules, reported progressive decrease in subjective alertness across the duration of the study. Moreover, several aspects of job performance, including safety sensitive tasks, degraded during the same time period. This suggests that current Federal regulations governing Hours of Service have the potential to allow work schedules which degrade the job performance of locomotive engineers and reduce the safety of railroad operations.					
17. Key Words locomotive engineers, fatigue, work schedules, hours of service, sleep, RALES simulator, work performance, train-handling, alertness; circadian rhythms, safety			18. Distribution Statement This document is available to the public through the National Technical Information Services, Springfield, VA 22161		
19. Security Classif. (of this report) UNCLASSIFIED		20. Security Classif. (of this page) UNCLASSIFIED		21. No. of Pages 34	22. Price

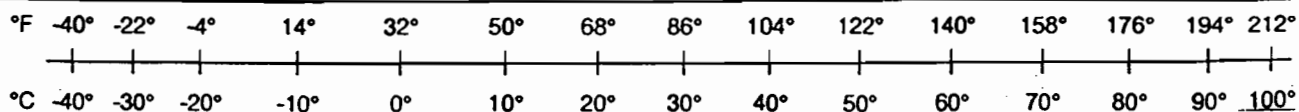
METRIC/ENGLISH CONVERSION FACTORS

ENGLISH TO METRIC	METRIC TO ENGLISH
LENGTH (APPROXIMATE) 1 inch (in) = 2.5 centimeters (cm) 1 foot (ft) = 30 centimeters (cm) 1 yard (yd) = 0.9 meter (m) 1 mile (mi) = 1.6 kilometers (km)	LENGTH (APPROXIMATE) 1 millimeter (mm) = 0.04 inch (in) 1 centimeter (cm) = 0.4 inch (in) 1 meter (m) = 3.3 feet (ft) 1 meter (m) = 1.1 yards (yd) 1 kilometer (km) = 0.6 mile (mi)
AREA (APPROXIMATE) 1 square inch (sq in, in ²) = 6.5 square centimeters (cm ²) 1 square foot (sq ft, ft ²) = 0.09 square meter (m ²) 1 square yard (sq yd, yd ²) = 0.8 square meter (m ²) 1 square mile (sq mi, mi ²) = 2.6 square kilometers (km ²) 1 acre = 0.4 hectare (ha) = 4,000 square meters (m ²)	AREA (APPROXIMATE) 1 square centimeter (cm ²) = 0.16 square inch (sq in, in ²) 1 square meter (m ²) = 1.2 square yards (sq yd, yd ²) 1 square kilometer (km ²) = 0.4 square mile (sq mi, mi ²) 10,000 square meters (m ²) = 1 hectare (ha) = 2.5 acres
MASS - WEIGHT (APPROXIMATE) 1 ounce (oz) = 28 grams (gm) 1 pound (lb) = .45 kilogram (kg) 1 short ton = 2,000 pounds (lb) = 0.9 tonne (t)	MASS - WEIGHT (APPROXIMATE) 1 gram (gm) = 0.036 ounce (oz) 1 kilogram (kg) = 2.2 pounds (lb) 1 tonne (t) = 1,000 kilograms (kg) = 1.1 short tons
VOLUME (APPROXIMATE) 1 teaspoon (tsp) = 5 milliliters (ml) 1 tablespoon (tbsp) = 15 milliliters (ml) 1 fluid ounce (fl oz) = 30 milliliters (ml) 1 cup (c) = 0.24 liter (l) 1 pint (pt) = 0.47 liter (l) 1 quart (qt) = 0.96 liter (l) 1 gallon (gal) = 3.8 liters (l) 1 cubic foot (cu ft, ft ³) = 0.03 cubic meter (m ³) 1 cubic yard (cu yd, yd ³) = 0.76 cubic meter (m ³)	VOLUME (APPROXIMATE) 1 milliliter (ml) = 0.03 fluid ounce (fl oz) 1 liter (l) = 2.1 pints (pt) 1 liter (l) = 1.06 quarts (qt) 1 liter (l) = 0.26 gallon (gal) 1 cubic meter (m ³) = 36 cubic feet (cu ft, ft ³) 1 cubic meter (m ³) = 1.3 cubic yards (cu yd, yd ³)
TEMPERATURE (EXACT) $[(x - 32)(5/9)]^{\circ}\text{F} = y^{\circ}\text{C}$	TEMPERATURE (EXACT) $[(9/5)(y + 32)]^{\circ}\text{C} = x^{\circ}\text{F}$

QUICK INCH-CENTIMETER LENGTH CONVERSION



QUICK FAHRENHEIT-CELSIUS TEMPERATURE CONVERSION



For more exact and or other conversion factors, see NIST Miscellaneous Publication 286, Units of Weights and Measures. Price \$2.50. SD Catalog No. C13 10286.

Updated 9/29/95

TABLE OF CONTENTS

<i>Section</i>	<i>Page</i>
Introduction	1
Method	3
Participating Engineers.....	3
The Simulator.....	3
Procedure.....	4
General.....	4
The Run.....	5
The Shift Schedules.....	6
The fast, backward rotating schedule.....	6
The slow, backward rotating schedule.....	7
Data Analysis.....	7
Performance measures.....	8
Results	8
Age, Experience, Schedule, Sleep and Time of Work.....	8
Sleep Debt, Time of Day and Alertness.....	10
Global Vigor.....	11
Performance measures.....	12
Analysis.....	12
Missed horns.....	13
Alerter response times.....	16
Fuel consumption.....	16
Discussion	18
Conclusions	19
References	20
Author Notes	22
Footnotes	23
Appendix I	28

LIST OF ILLUSTRATIONS

<i>Figure</i>	<i>Page</i>
1. Model of the effect of work schedule on job performance.....	3
2. RALES viewed from inside (left) and outside (right).....	5
3. Segments within each run.....	9
4. Actigraph measures of sleep duration in minutes as a function of runs for each group.....	10
5. Actigraph measures of sleep onset in minutes as a function of runs for each group....	12
6. Predicted alertness for the Fast and Low Groups as a function of segments. Each run has ten segments.....	13
7. Mean Global Vigor scores and predicted alertness for the Fast Group as a function of time of day across runs 3, 4, 5 and 6.....	14
8. Mean Global Vigor scores and predicted alertness for the Slow Group as a function of time of day across runs 3, 4, 5 and 6.....	14
9. Missed horns for the Fast and Slow Groups as a function of run.....	15
10. Mean alert latency for the Fast and Slow Groups as a function of run.....	15
11. Mean fuel consumption for the Fast and Slow Groups as a function of run.....	17
12. Mean cumulative deviation in fuel consumption from run 3 for the Fast and Slow Groups as a function of segment.....	17

LIST OF TABLES

<i>Table</i>		<i>Page</i>
1.	Shift Schedule - Fast Backwards Rotating Group ("Fast Group").....	6
2.	Shift Schedule - Slow Backwards Rotating Group ("Slow Group").....	7
3.	Summary of Sleep and Work/Rest Schedule Characteristics.....	11

EXECUTIVE SUMMARY

Under current law, Federal regulations, and railroad crew-calling practices, a locomotive engineer can work for many consecutive days, at varying times of day, and for durations of up to 12 hours per shift. Consequently, a locomotive engineer's work schedule might never develop a consistent circadian pattern. This study examines the impact of an example of such practices on locomotive engineer performance.

To investigate the effect of work schedules on train handling performance and vigilance, certified locomotive engineers performed normal job duties while operating a highly realistic locomotive simulator on two different work schedules. The work schedules were designed to conform with Federal regulations governing Hours of Service, to cause shift start times to rotate backwards (shifts start earlier each day) at different rates, and to produce different levels of sleep deprivation. Both schedules required ten hours of work, but differed in the amount of off-duty time.

The faster backward rotating schedule caused shift start times to begin four hours earlier each day and resulted in an average sleep duration of 4.6 hours. The slower backward rotating schedule caused shift start times to begin two hours earlier each day and resulted in an average sleep duration of 6.1 hours. Subjective measures of alertness (Global Vigor Scale) under both schedules declined over the course of each work period and across successive work periods.

Job performance was continuously monitored during the course of each work day. Alertness declines were observed in both engineer groups. Failures to sound the train horn at grade crossings, response times to the audible warning on the alerter, and fuel use were observed to increase across successive work periods in both groups.

INTRODUCTION

As a result of current law ¹, Federal regulations ², and railroad crew-calling practices ³, a locomotive engineer can work for many consecutive days, at widely varying times of day, and for durations of 12 hours per shift. The Hours of Service Act of 1907, as amended, stipulates that locomotive engineers may work no more than 12 continuous hours without a minimum of 10 hours off duty, and that they be given at least 8 consecutive hours off duty in every 24-hour period. An individual can work 11 hours and 59 minutes, be off duty for 8 hours, and return to work at the end of that 8-hour period. Moreover, such a pattern could continue for many consecutive days ⁴, so that the individual's work schedule would never develop a consistent circadian pattern. Crew members are generally called approximately 2 hours before reporting time, so that the maximum duration of uninterrupted sleep could be between 6 and 7 hours. However, since the required 8 hours off-duty time includes travel, leisure and personal time, the duration of any sleep would be even less than that.

Do current practices impose an excessive burden of sleep deprivation, circadian disruption, and fatigue which could degrade the train handling performance and vigilance of locomotive engineers?⁵ There are several sources of information, scientific and anecdotal, which indicate that while the average locomotive engineer does not have a work schedule consisting of many days of an alternating 11:59 on, 8:00 off cycle, the average schedule *is* conducive to fatigue and circadian disruption ⁶. For instance, a 1992 GAO study of 124 engineers found that the average shift duration was approximately 8 hours. However, nearly 19% of the engineers studied had work periods between 8 and 10 hours in duration and nearly 15% had periods between 10 and 12 hours in duration. Since the average shift start time variability was approximately 4 hours, many engineers work long hours at irregular times. A direct result of long, irregular work hours is often inadequate sleep time.

This is supported by a recent study of work-rest schedules (see footnote 6, Pollard, 1996), which found that nearly 25% of surveyed locomotive engineers slept 6 or fewer hours per night. Moreover, inadequate rest time because of long work hours has been a complaint of labor unions. In Senate testimony in 1991, Larry D. McFather, President of the Brotherhood of Locomotive Engineers, stated: "Another of the safety hazards is that crews are not given 10 hours off duty for rest as provided by 45 U.S.C. § 62(a) (1). The reason is that the railroads require an employee to mark off at 11 hours and 59 minutes, even though he/she actually remains on the train much longer and is not free to leave." The National Transportation Safety Board has documented cases in which engineer fatigue, caused by inadequate sleep and a long, irregular work schedule, has resulted in serious accidents. The seriousness of this situation within the industry is indicated by the recent formation of a labor-management task force to study work-rest schedules, as noted in the 1994 Report from the Committee on Energy and Commerce on H.R. 4545 to amend the Federal Railroad Safety Act of 1970.

The research reported here describes an approach to determining the potential effect of the work schedules of locomotive engineers on train handling performance and vigilance through the use of two work schedules, each one week long, which comply with The Hours of Service Act, and the Research and Locomotive Evaluator/Simulator (RALES) simulator. The work schedules and the RALES simulator are designed to allow an examination of job performance within the context of the work week schedules that a locomotive engineer might encounter. Our approach is displayed graphically in Fig. 1. In this model, work schedules affect the time of day of work and a variety of sleep characteristics. Time of work and sleep, in turn, affect alertness, which in turn affects work performance.

The rationale of this approach is motivated by the relatively large number of interconnected variables which must be considered. The principal variables of concern in studying the effects of work schedules on job performance are the duration of work, the time of day of work, and the variability of the start of work periods. Obviously, these variables are not unrelated, so that manipulations of one variable can affect the others. However, only a systematic variation of each variable within the constraints of a natural 24-hour day and the limitations of the Hours of Service Act can adequately discern the effect of each variable on job performance. While these variables are important determinants of performance in other occupations and on other work schedules⁷, there is no direct evidence to suggest that this is also the case for locomotive engineers. In addition to the schedule variables already noted (duration of work, time of day of work, and start time variability), there are a number of variables which are inherent to, and/or controlled by, the subject. These variables are partially controlled by the work schedule and may be extremely important for job performance. These include sleep duration, quality of sleep, hours since last consolidated sleep, and cumulative sleep deprivation over the week. However, the demand characteristics of the engineer's job and self-selection for this type of work schedule could mitigate any detrimental effects which had been seen in a different context. Thus a full-scale experimental investigation of this problem is risky.

The alternative is to pursue a more limited objective: namely, the comparison of within-subject performances on work schedules which comply with the Hours of Service Act, but which differentially affect sleep and work characteristics. Two schedules were devised which had the same total duration of work over the course of the test period, but which differed in the start times of the work and the rest time available for sleep between work periods. The time of work and accumulated sleep deprivation (sleep debt) on these schedules were hypothesized to differentially affect alertness and job performance. The sleep characteristics of the locomotive engineers was obtained by the use of an activity measurement technique called actigraphy which allowed the status of these variables to be determined for the period of time prior to each experimental run. Alertness was predicted from time of day and sleep debt as determined by actigraphy measures.



Figure 1. Model of the effect of work schedule on job performance

METHOD

Participating Engineers

Fifty-five certified locomotive engineers⁸ were recruited from the local area to participate. Each engineer was paid \$250 per day. There was a \$50 penalty imposed for violations of speed limits or missed stops. Each engineer was given a brief description of the general purpose of the study and was told that their hours of work during the study would conform with the Hours of Service Act.

The Simulator

Participating engineers performed their work duties by operating FRA's Research and Locomotive Evaluator/Simulator (RALES) at the IIT Research Institute (IITRI) in Chicago. The RALES unit consists of a fully equipped locomotive cab which is mounted on a six-axis motion base to duplicate a locomotive's motion. The hydraulically actuated motion base is controlled by a computer model of the track geometry and locomotive suspension characteristics and is used to cause the cab to move as if it were really proceeding down the track. Motions include bounce and rock associated with switches, bridges, road crossings and rail joints, lateral motion within the gauge and in curves, longitudinal impacts resulting from coupler slack run-in and run-out, and steady state acceleration and deceleration during motoring and braking. During simulation runs, synchronized laser disk projection shows a wide-screen forward, side and ballast view. Transitions between day and night are visually realistic and appropriate. The RALES simulator is provided with changing day/night video projection capabilities so that the visual display always corresponds with the actual time of day outdoors. As real world sunset or sunrise light levels change, the video projection keeps track. These changes are done with video special effects manipulations which include a field-of-

view at night dominated by simulated headlamp illumination. In addition, sounds of the locomotive environment are also simulated. This includes blowing the horn, operation of way-side crossing gate bells (including Doppler-effect), air brake pipe exhaust and engine sounds which vary as throttle position and engine load are varied. Figure 2 shows two views of RALES.

RALES is computer controlled so as to simulate changes in train handling which occur as a result of increases in throttle, application of different braking systems (independent, automatic, and dynamic), train composition (load, number and types of cars, amount of motive power, distribution of weight), and terrain (undulating, graded, etc). As a result, it provides an accurate and realistic simulation of train behavior in response to the engineer's actions and omissions. Changes in speed, brake pipe pressure, within-train forces, slack, and consumption of fuel are all accurately responsive to the engineer's manipulation of the locomotive controls. The realism of the simulation is such that 85% of a group of 53 professional railroad engineers indicated that there were times while using RALES that they forgot that they weren't operating a real locomotive.

Two trains were simulated in the experiment. The first train, used only for familiarization runs on the first day⁹, was a "hot" intermodal train with the characteristics of high power to weight ratio and relatively easy handling. Train orders for this train were prepared to allow each engineer to make two runs over a territory of 190 miles in one eight-hour shift.

The second train was a unit coal train of 88 cars and four locomotives. The train was 5,000 feet in length and had a gross weight of 12,400 tons. Total power was 12,000 HP for a total of 0.98 HP per ton. There were 134.6 tons per operative brake.¹⁰

The locomotive cab temperature was maintained at 78° F. Real time VHF railroad radio audio was added to the cab and locomotive sound was maintained at real world levels. The alerter¹¹ was functional, with light flash followed after an interval by sound alarm. Failure to respond to the sound alarm, after an interval, would result in an emergency application of the air brakes.

During the study, engineers wore a wrist-borne activity monitor called an Actigraph¹². The device was worn on the left wrist and was removed only for showers. During shifts the simulator cab movement was the dominant motion recorded by the device. Off-duty period data, however, reflected engineer activity, and these periods were assessed with Action 3 personal computer software¹³ to determine sleep during the off-duty period, total number of awakenings, and latency to sleep onset.

Procedure

General

During runs, the RALES operator played the role of dispatcher, and an in-cab observer



↑ Locomotive Engineer operating RALES ⇒

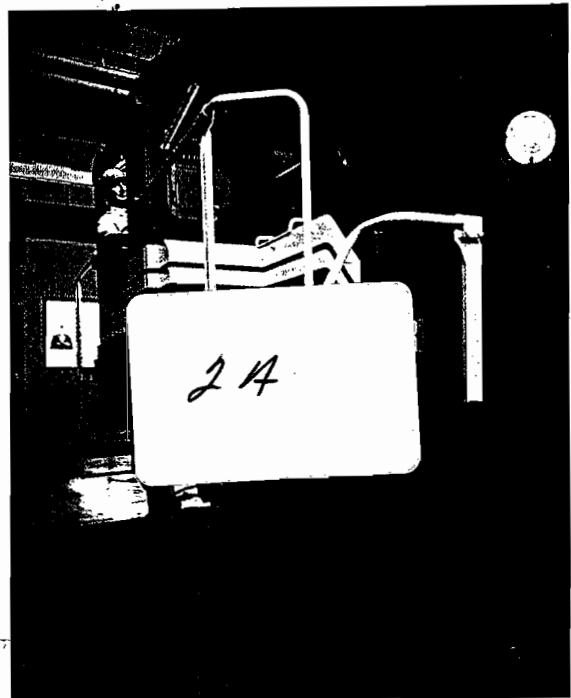


Figure 2. RALES viewed from inside (left) and outside (right).

played the role of conductor. During runs, infrared light source video tape records were made of each engineer.

Runs and shift schedules were prepared to represent, as realistically as possible, the conditions encountered by over-the-road engineers in terms of tasks, shift rotations, and shift durations. The schedule and activities chosen were ones which were likely to provoke fatigue and decrease vigilance.

The Run

The run was 190 miles long and was used for each work shift. The run covered terrain which is representative of the Midwest United States. This terrain lacks the extremes of grade and curvature found in other locations and is, at numerous locations, typified by the Association of American Railroads' designation "undulating terrain." This territory was felt to be within the capabilities and experience of the engineers.

The engineers were provided with a paper timetable and a track profile. These resources are generally available to locomotive engineers. Engineers were expected to learn and navigate the territory on this basis. One concession made in regard to identifying track location was the provision of a milepost video monitor display in the locomotive cab. The monitor would display milepost

number at a location point where the milepost could be seen in the real world and continue to display that number until the milepost location was past. Hence, the milepost number was only visible for as long as it would be normally. This alteration was necessary due to the difficulty in achieving a sufficient level of dynamic light range in projected video images.

The runs were, from Tuesday morning onward, 10 hours in length. The run was divided into ten continuous segments, each with a red stop signal at its end. The run was managed by holding engineers at signals for an amount of time intended to pace time. Operations on each segment began at the same approximate time for all engineers making that run.

Actual stopping time averages ran from 15 to 30 minutes, depending on the running time the engineer had for the previous segment. Engineers were free to leave the locomotive cab during stops, but were not told the intended duration of the individual stop.

At each of the ten stops during a run, engineers responded to a computerized visual analogue scale to measure general subjective level of vigor and affective state ¹⁴. The Global Vigor scale assesses subjective alertness, sleepiness, effort and weariness. The Global Affect scale assesses subjective emotional state (happy, calm, sad, tense).

The engineers were advised that they would have to provide their own food, just as they would in their daily work. A 24-hour convenience store was located within walking distance of the place of rest. Cold water was provided aboard the locomotive. Personal rest stops during the experimental runs were limited to stops at red signals (there were nine such stops during a run, plus a final red signal at the end). There were no other stops or breaks.

The Shift Schedules

Table 1.
Shift Schedule - Fast Backwards Rotating Group ("Fast Group").

DAY	HOURS	RUNS
Monday	8:00 am to 6:00 pm	1 & 2
Tuesday	8:00 am to 6:00 pm	3
Wednesday	2:00 am to Noon	4
Wednesday	10:00 pm to 8:00 am Thursday	5
Thursday	4:00 pm to 2:00 am Friday	6

The fast backward rotating schedule. The fast backward rotating schedule was arranged so as to agree with hours of service rules, but also to provide as disruptive an effect as possible on

normal sleep cycles. The schedule began on a Monday. In all cases, the engineers did not work the prior two days. The Monday shift and the Tuesday morning shift were arranged to encourage adaptation to day work. Following the 10-hour run beginning Tuesday morning, however, the schedule began to follow hours of service rules with shift times which rolled backwards, causing the engineers to arise earlier each day. Table 1 presents the fast schedule. The engineers were not told what the work schedule would be, although by the end of Run 4, they all independently stated that they assumed it would continue to follow Hours of Service rules.

The place of rest, a motel, was located approximately 20 minutes from the simulator site with transportation provided. Crew calls, as is customary, were made two hours prior to the beginning of the shift. As an additional disruption of sleep, engineers were always given a "wrong number" telephone call at the midpoint of their rest period, beginning on Tuesday evening.

During fast schedule runs, the dispatcher was instructed not to respond to engineer calls and to limit communications to those indicated by the experiment script. The conductor was instructed to call signals¹⁵ only when the engineer did and to converse only in response to engineer initiated discussion. There were 35 engineers in the Fast Group.

The slow backward rotating schedule. The slow backward rotating schedule was identical to the fast schedule on Monday and Tuesday. The following days, the slow schedule required 10-hour work shifts as on the fast schedule, but also allowed a longer off-duty period between shifts. As a result, the slow schedule, which is presented in Table 2, also rotated backwards over days but at a slower rate. During slow runs, the conductor was instructed to call signals and to initiate conversation with the engineer at 15-minute intervals. Otherwise, all other conditions remained the same as for fast runs. There were 20 engineers in the Slow Group.

Table 2.
Shift Schedule - Slow Backwards Rotating Group ("Slow Group").

DAY	HOURS	RUN
Monday	8:00 am to 6:00 pm	1 & 2
Tuesday	8:00 am to 6:00 pm	3
Wednesday	6:00 am to 4:00 pm	4
Thursday	4:00 am to 2:00 pm	5
Friday	2:00 am to Noon	6

Data Analysis

The run was divided, for purposes of schedule control and subsequent data analysis, into ten

segments of roughly equivalent time durations. This is schematically presented in Figure 3. The segments ran from stop signal to stop signal to afford "dispatcher" control over the run. This control allowed the shift length to be controlled by manipulating the delays at stop signals while holding the running conditions of each segment constant regardless of shift length. Run scheduling was organized in such a way as to cause all engineers to run at the same clock time over each segment for a given scheduled shift. This organization afforded a comparison of train handling across engineers for specific tasks at specific times. An additional four noncontiguous segments (labeled csp 1, etc. in Fig. 3) were superimposed on the run to assess the ability of the engineers to maintain a constant speed over varied terrain.

Performance measures. A wide variety of job-specific performance measures were obtained in this study. However, for the sake of brevity, this report focuses on only those measures which appear to be particularly pertinent to safety and operational efficiency issues. Additional discussion of other performance measures is available in a technical report ¹⁶.

Two of the performance measures were simple response items: blowing the horn for grade crossings and responding promptly to alerter flashes. Most grade crossings during the run were marked with whistle boards ¹⁷. The number of times an individual failed to blow the horn for a crossing within each of the ten stop-to-stop segments was counted. Latencies of responses to the alerter were recorded as the alerter progressed from flashing to an auditory signal.

The expenditure of fuel by the engineer within each of the stop-to-stop segments was also measured. Fuel was measured in gallons and was calculated from the dynamic computer model of the train. A high expenditure of fuel by the engineer indicates that the engineer is aggressively accelerating and braking the train, or powering the train while the brakes are applied. Engineers are trained and constantly reminded to avoid this type of operation since it increases fuel consumption and can be the cause of excessive slack action, which is a primary cause of engineer-caused accidents. An increase in fuel expenditure by a given engineer is likely to be caused by a lack of attention to train control. For instance, delayed reactions to changes in train speed would result in heavier use of throttle or braking systems in order to maintain control.

RESULTS

Age, Experience, Schedule, Sleep and Time of Work

The Fast Group engineers had an average age of 43.3 years as compared with the Slow Group engineers whose average age was 40.5 years. Average years of service for the two groups was 14.5 and 13.85, respectively. Neither factor was reliably different between groups ¹⁸

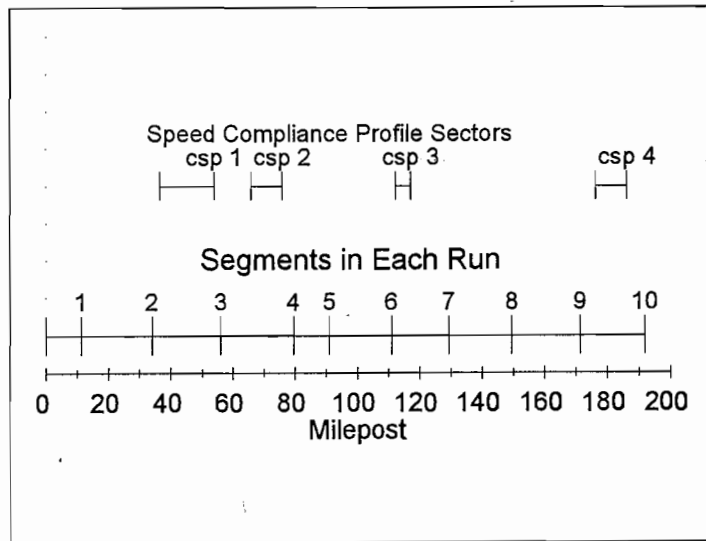


Figure 3. Segments within each run

The two groups were exposed to work schedules which were both designed to produce progressive sleep loss over the course of the experiment. Moreover, it was expected that sleep loss would be greater for the Fast Group than for the Slow Group. The actigraph measures of sleep duration as a function of runs are presented in Figure 4 for each group. Both groups began run 3 with approximately the same amount of sleep, but the Fast Group got less sleep than the Slow Group on all subsequent runs. Both groups exhibited a decrease in sleep durations relative to run 3 on subsequent runs, although the pattern differed between groups. The Fast Group had the least sleep before run 4, while the Slow Group had the least sleep before run 6. A nonparametric analysis of variance¹⁹ indicated that there were reliable differences between groups and between runs, as well as a reliable group by runs interaction²⁰.

Figure 5 shows sleep onset as a function of runs for each group. Sleep onset is similar to the multiple sleep latency measure²¹ in that nonsleep deprived individuals have long latencies to sleep, while sleep deprivation progressively decreases sleep latency. As can be seen in Fig. 5, both groups have approximately the same latency before run 3. However, in subsequent runs latencies become shorter, with each group showing a different pattern. A nonparametric analysis of variance again showed reliable group, run, and interaction effects²².

The two schedules were also designed to differentially affect time of work in the two groups. For the Fast Group, work start times began four hours earlier on each successive shift. By comparison, work start times began two hours earlier on each successive shift for the Slow Group. Table 3 summarizes important sleep and work schedule characteristics for the two groups.

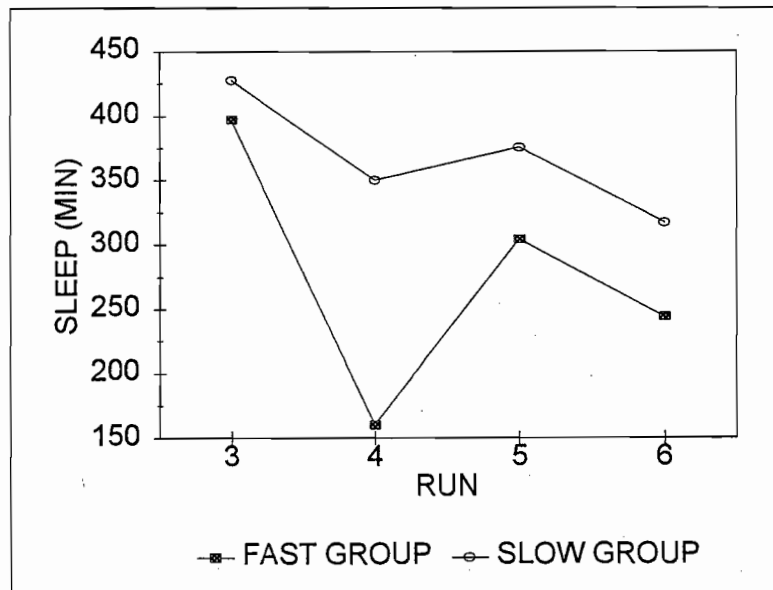


Figure 4. Actigraph measures of sleep duration in minutes as a function of runs for each group.

Sleep Debt, Time Of Day And Alertness

The pattern of amount of sleep and sleep onset for each group suggests that the two work schedules had the desired effect of producing a progressive sleep debt in each group which differed in rate of accumulation and final, total sleep debt. Sleep debt and time of day are two of the most important variables that modulate job performance. Numerous studies have demonstrated that, regardless of prior sleep, performance is at a nadir in the vicinity of 3 A.M.²³ Time of day is critically important in the present context because each schedule rotates backward, with the result that the segment for which the poorest performance is expected would change on each run. Moreover, the 10 segments in each run differ in their performance demands in keeping with the goal of the study to maintain a maximum level of realism. Consequently, the subsequent analyses below compare individual segments across runs with respect to sleep debt and circadian time of day. An alertness model, described below, is used for this purpose.

As indicated in Appendix I, time of day influences on alertness can be modeled by a simple cosine function with a period of 24 hours. An "offset" is added to this function so that the output of the model is always a positive number. The offset is an exponential decay function of total sleep debt, so that the offset has a range between 1 and 0. Sleep debt accumulates at a rate of 1 hour for each hour awake, and it is repaid at a rate of 1 hour for each hour asleep. Predicted alertness is the output of the model which is 100 times the offset plus the cosine function. A comparison of the predicted alertness on the two schedules as a function of segment is shown in Fig. 6. Note that since each run has 10 segments, it is easy to visually distinguish the daily runs.

Table 3.
Summary of Sleep and Work/Rest Schedule Characteristics.

	WORK SCHEDULE	
	FAST	SLOW
SLEEP DURATION	4.6 HOURS	6.1 HOURS
SLEEP LATENCY	2.3 HOURS	3.0 HOURS
START TIME DIFFERENCE	- 4 HOURS	-2 HOURS
HOURS OFF-DUTY	9.3 HOURS	12 HOURS
SHIFT LENGTH	10 HOURS	10 HOURS

The predicted patterns of alertness for the two groups are identical in run 3. Alertness is predicted to differ to different degrees on subsequent runs and depends critically on the specific segment which is examined. For example, the first segment in the Fast Group decreases between runs 3 and 4 but increases on runs 5 and 6. In the Slow Group, there is a steady decline in alertness across runs for this same segment. The last segment shows a different pattern. For the Fast Group predicted alertness declines from run 3 to 4, increases from 4 to 5, and then decreases to run 6. For the Slow Group, there are slight increases from run 3 to 5 followed by a slight decrease on run 6. It is clear that the pattern of predicted alertness varies between and within groups in a complex fashion.

Global Vigor. The Global Vigor scale measures subjective fatigue and alertness and should, therefore, be sensitive to the combined effects of sleep debt and circadian time of day. Since it is not a performance measure, it should not be sensitive to skill or other differences between the groups or between segments. Therefore, on the basis of the alertness model, global vigor is not expected to differ between groups on run 3 because sleep debt and circadian time of day are similar for the two groups in run 3.

Figures 7 and 8 show the mean Global Vigor scale values and the alertness predictions for the Fast and Slow Groups. The alertness model shows good agreement with global vigor across groups and segments²⁴. There is a clear pattern for both groups. The model and scale values do not differ between groups during run 3. Subjective ratings of alertness and lack of fatigue are highest at the beginning of each run and decrease progressively to the end of the run. The groups differ with respect to changes in global vigor across runs, however. The Fast Group shows a decrease between run 3 and 4 and slight increases from runs 4 to 6. The Slow Group, on the other hand, shows a progressive decline across runs. A nonparametric analysis of variance for Global

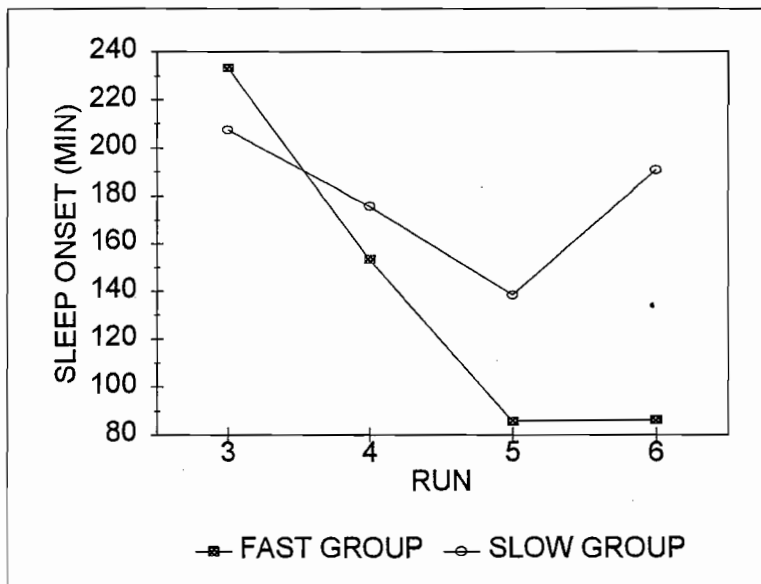


Figure 5. Acitgraph measures of sleep onset in minutes as a function of runs for each group.

Vigor confirmed these observations. It showed reliable segment, runs, and runs by group effects ²⁵.

Performance measures.

Analysis. From the previous section, it should be clear that performance measures in this experiment are not capable of a simple analysis. There are two reasons for this. First, after run 3, the same segment in each run occurs at a different time of day within and between groups. Consequently, the sleep debt in each segment also differs within and between groups. If alertness (time of day and sleep debt) affects performance, then performance will be modulated within and between groups in the pattern depicted in Fig. 6. Second, segments differ considerably from one another with regard to operating requirements demanded of the locomotive engineer. Segments each have various combinations of mileage, curvature, grade, speed limits, grade crossings, signals, etc. Every performance measure can be expected to vary across segments because of the differences in operating requirements between segments.

Because segments differ, only performances across runs on the same segment should be compared. Because the same segment in each run occurs at different times of day between and within groups, comparisons of performance must take relative sleep debt and time of day influences into account. The alertness model is used in the following analysis of performance measures because it allows individual segments to be compared on the basis of relative sleep debt and

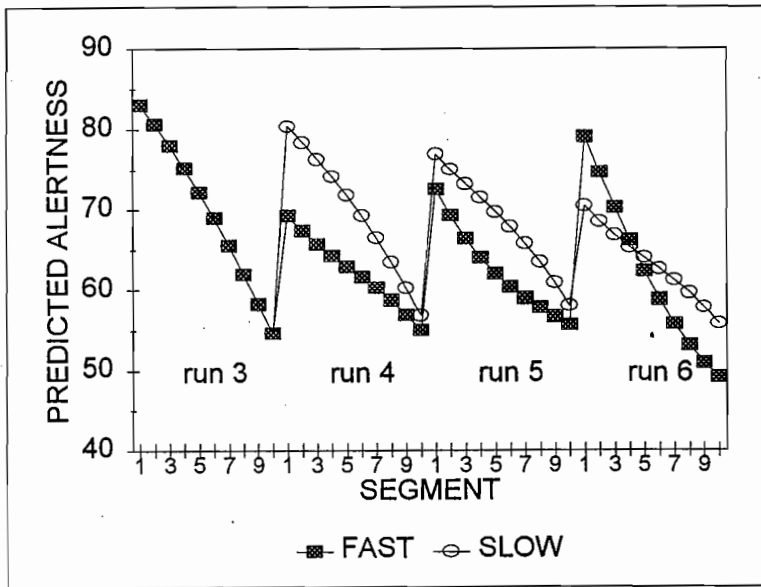


Figure 6. Predicted alertness for the Fast and Slow Groups as a function of segments. Each run has ten segments.

circadian time of day (i.e., predicted alertness). Without the alertness model the effects of these variables are confounded.

The alertness model suggests that, on the basis of sleep debt and circadian time of day alone, there should be no differences in performance between groups on run 3. Global Vigor, which is not a performance measure but a measure of fatigue and alertness, validates this prediction of the model. A group difference in a performance measure on run 3 would indicate that performance was affected by some factor other than the sleep debt and circadian time of day influences (alertness) operating under the two schedules. Since the alertness model does not consider group differences in skill or other factors, the model will be inadequate when such differences exist. A strategy was adopted, therefore, for the analysis of performance measures in which it was first determined if the groups differed during run 3. In the absence of difference, analysis would proceed as for the global vigor scale data. Otherwise, each group's performance would be analyzed using its performance on run 3 as a baseline.

Missed horns. Figure 9 shows the mean number of missed horns for the Fast and Slow Groups as a function of run. Missed horns do not differ statistically between groups on run 3. A non-parametric analysis of variance further indicated that there were no overall group or group by run effects, although there was a reliable runs effect. Since there were no reliable performance differences evident in run 3, the alertness model was used to reanalyze the pattern of missed horns with regard to sleep debt and circadian time of day. A reliable negative correlation was found between missed horns and changes in performance predicted by the model. The negative correlation

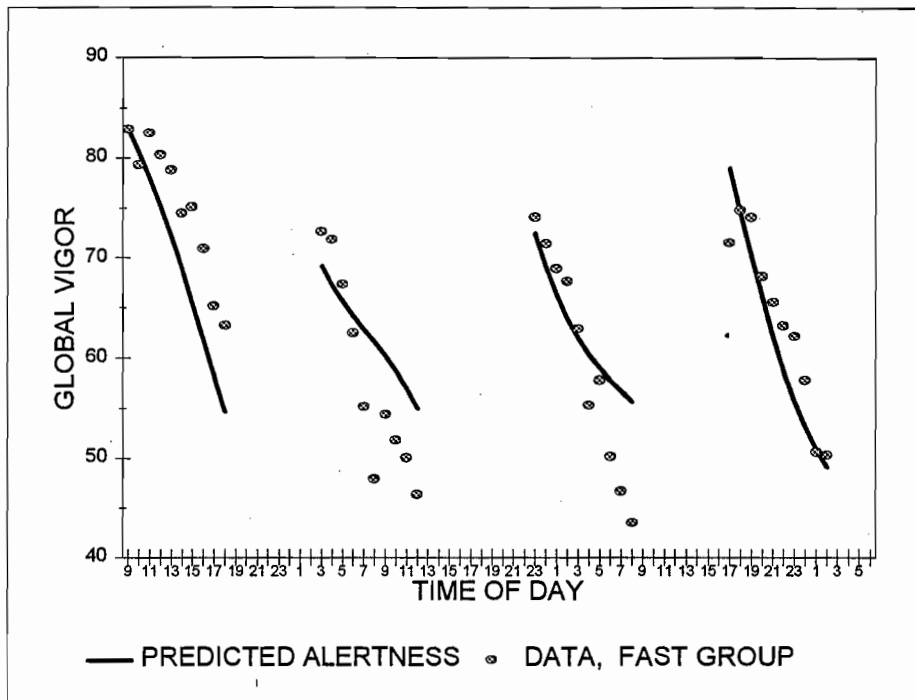


Figure 7. Mean Global Vigor scores and predicted alertness for the Fast Group as a function of time of day across runs 3, 4, 5 and 6.

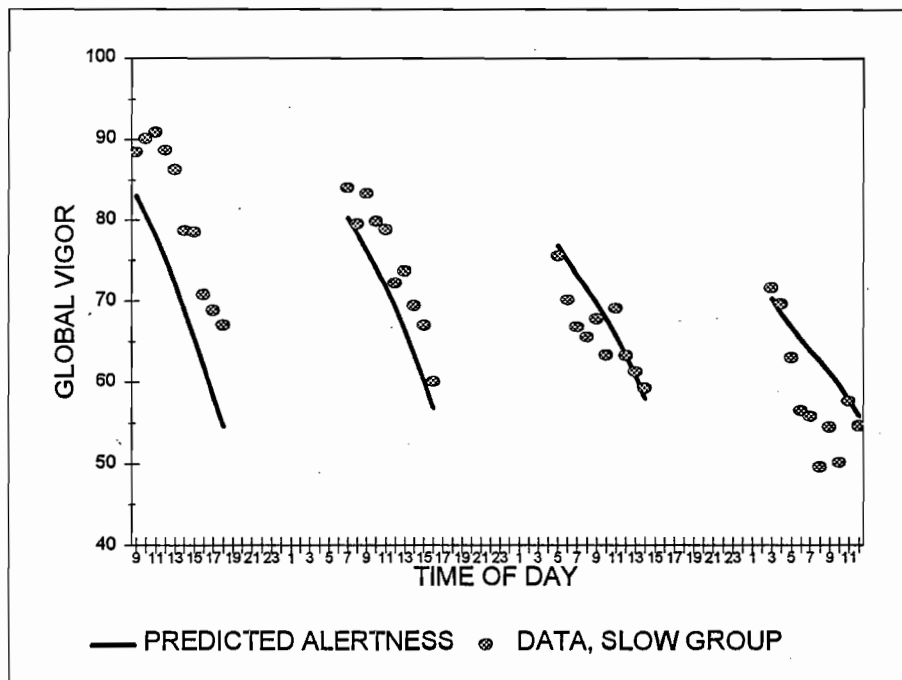


Figure 8. Mean Global Vigor scores and predicted alertness for the Slow Group as a function of time of day across runs 3, 4, 5, and 6.

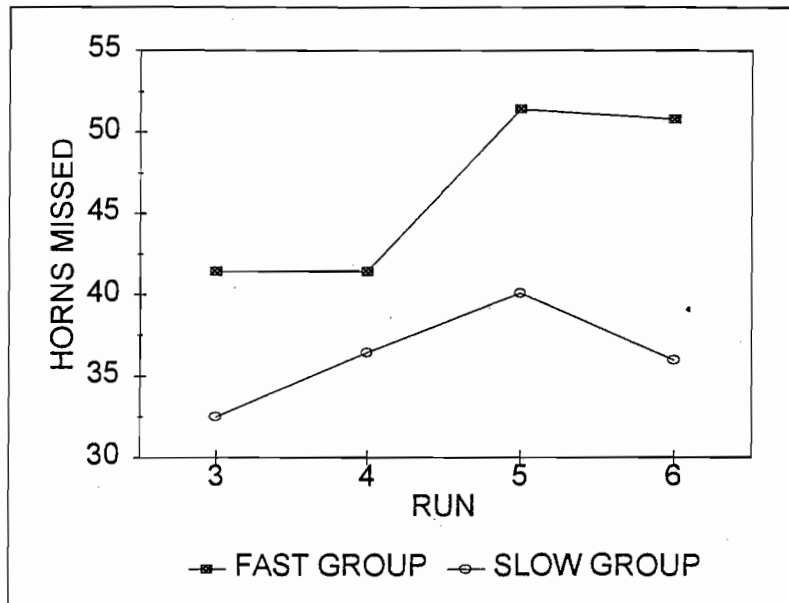


Figure 9. Missed horns for the Fast and Slow Groups as a function of run.

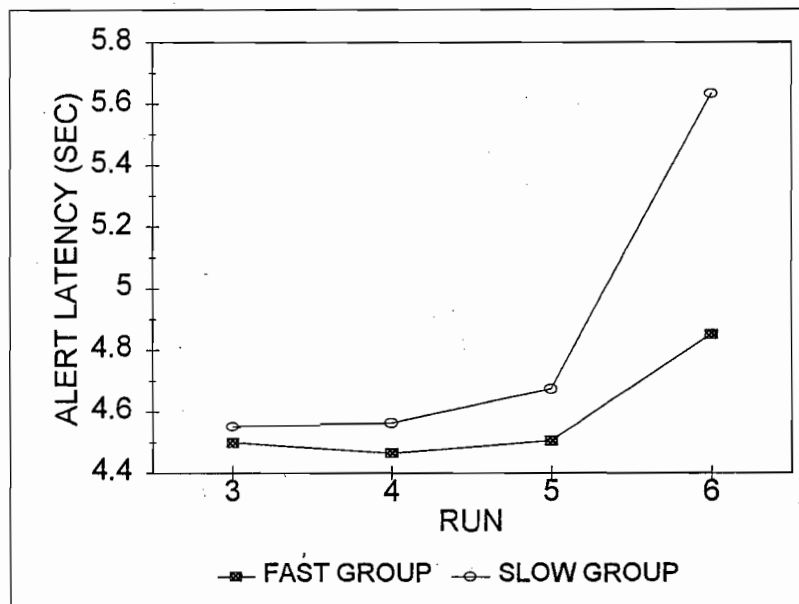


Figure 10. Mean alert latency for the Fast and Slow Groups as a function of run.

indicates that more horns are missed as performance is predicted to decline ²⁶.

Alerter response times. Figure 10 shows the mean alerter response times for the Fast and Slow Groups as a function of runs. Again, alerter response times do not differ statistically between groups for run 3. However, a non-parametric analysis of variance indicated reliable group, runs and group by runs effects. In the absence of reliable run 3 differences between the groups, the alertness model was used to analyze the alerter response times. In this instance, however, the correlation was not reliable, although it was in the right direction ²⁷.

Fuel consumption. Figure 11 shows the mean fuel consumption for the Fast and Slow Groups as a function of runs. Fuel consumption was higher in the Slow Group than in the Fast Group in run 3. A non-parametric analysis of variance indicated overall group and runs effects, but no group by run interaction ²⁸. It is apparent that the Slow Group consumed more fuel on every run than the Fast Group, which is highly suggestive of a difference in operating skill.

Performance in runs 4, 5 and 6 was consequently rescaled for each group relative to its run 3 performance. Fuel use in each of the ten segments in run 3 was subtracted from the same segment in runs 4, 5 and 6. Improved performance on subsequent runs would then result in positive scores, while decrements in performance would result in negative scores. To facilitate comparison with the performance model, consecutive difference scores across segments and runs were cumulated. The resulting performance is shown in Fig. 12.

From Fig. 12 it is clear that performance decrements occurred for most segments in the Fast Group since the deviations from baseline have a continuous positive trend from run 4 through run 6. By contrast, the Slow Group does not exhibit a positive trend until run 5. This suggests that performance differences between groups was largest on run 4, which is consistent with the alertness predictions of the model (see Fig. 6). A reliable correlation was found between cumulative fuel deviations and predicted alertness ²⁹.

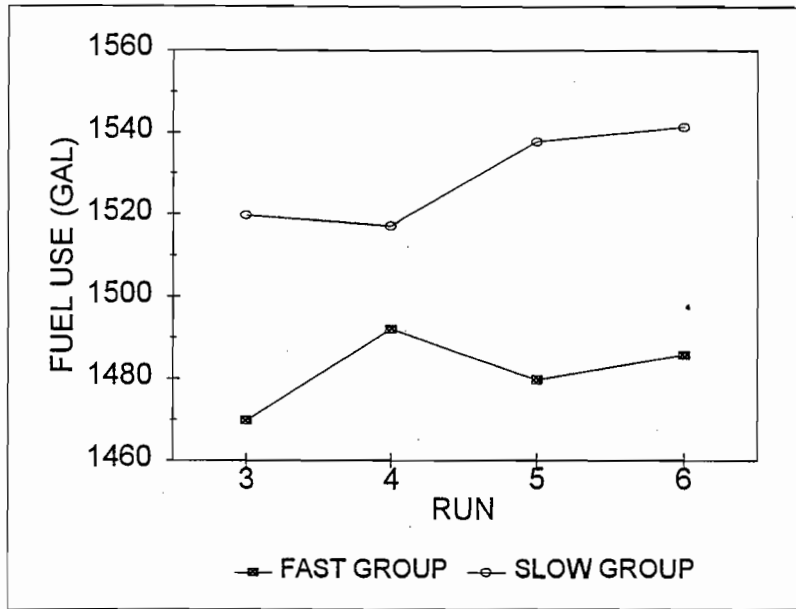


Figure 11. Mean fuel consumption for the Fast and Slow Groups as a function of run.

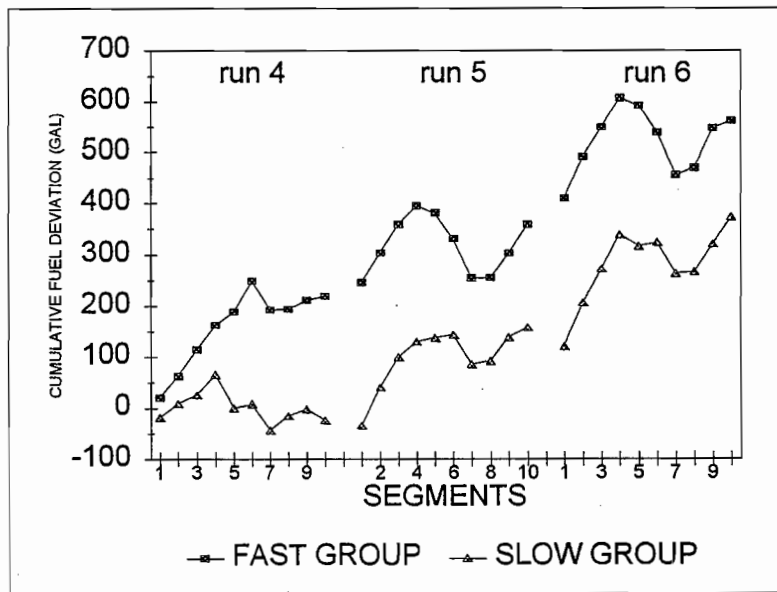


Figure 12. Mean cumulative deviation in fuel consumption from run 3 for the Fast and Slow Groups as a function of segment. The first 10 segments are not represented because these are the run 3 baseline.

DISCUSSION

The two work schedules described in this report are similar to work schedules that have been observed in railroad operations³⁰. It is not unusual for locomotive engineers to work for long periods of time on schedules in which the available time for sleep is less than 8 hours and includes travel time and time for personal hygiene and biological needs. The two work schedules were specifically designed to approximate that aspect of locomotive engineer schedules. Although both groups had identical shift lengths of 10 hours, the Fast Group averaged 9.3 hours off-duty while the Slow Group averaged 12 hours off-duty. It was anticipated that such schedules would, in turn, result in less sleep for both groups than the general population. While the general population obtains approximately 7.5 hours of sleep per night³¹, the Fast Group obtained 4.6 hours of sleep and the Slow Group obtained 6.1 hours of sleep. It is clear that the schedules had the desired characteristic of allowing both groups to accumulate a progressive sleep debt across the duration of the experiment, and that sleep debt was allowed to accumulate faster for the Fast Group.

However, individuals differ with respect to the amount of sleep per night that they require. The schedules in this experiment were designed to allow an accumulation of sleep debt relative to the amount of sleep required by the general population. It is possible that locomotive engineers self-select for a type of employment that requires long hours of work and less than 8 hours of nightly sleep because they require less than an average amount of sleep per night. Two pieces of information in this report indicate that the engineers were, indeed, sleep deprived during the course of the experiment. First, subjective ratings of alertness (Global Vigor, see Figs. 7 and 8) indicate that alertness declined over the course of each run in each group. After run 3, when the schedules began to restrict sleep duration, alertness was consistently lower in each segment of runs 4, 5 and 6 than for the same segment of run 3. This pattern of subjective alertness is consistent with an accumulation of sleep debt which is only partially repaid during intervening off-duty periods. If the engineers in this study were receiving as much sleep as they required, they would not have indicated a consistent pattern of lowered alertness on runs 4, 5 and 6. Second, the sleep onset measure in Table 3 and Fig. 5 indicates that engineers in the Fast Group went to sleep sooner than engineers in the Slow Group. A faster sleep onset is consistent with a faster accumulation of sleep debt in the Fast Group.

The two schedules were also designed to cause shifts to start earlier on each successive day. This backward rotation of shift start times is counter to the normal biological circadian rhythm which rotates slowly forward. The Fast Group rotated backwards 4 hours per day and the Slow Group rotated backwards 2 hours per day. This backward rotation was expected to exacerbate the effects of sleep deprivation on alertness and job performance. Alertness and job performance are often modeled, as in this report, as functions of sleep debt and time of day³². The alertness model in this report predicted subjective alertness based on accumulated sleep debt and time of day with 68% accuracy³³.

On the basis of the alertness model, it was expected that job performance would decline across runs in both groups. The data for missed horns, alerter response time and fuel consumption are consistent with the predicted pattern. For all three measures, performance declined from run 3 to run 6 as predicted. Both schedules were designed to allow the accumulation of a sleep debt across runs, the engineers subjective ratings of their own alertness confirmed that they felt progressively less alert across runs, and their job performance measures, correspondingly, declined across runs.

As noted previously, differences between the groups with regard to time of day at which particular segments occurred, differences between segments with regard to operating demands, and differences between the groups with regard to operating skill, all make a simple analysis of group differences in job performance extremely difficult. The alertness model was used as an analysis tool to overcome some of these difficulties. Reliable and similar correlations were obtained between predicted alertness and missed horns and cumulative deviations from baseline fuel consumption. These results are consistent with the expectation that individuals on more rapidly backward rotating schedules accumulate a sleep debt more rapidly and perform job tasks less efficiently and less safely.

The work schedules used in this study are not unusual for railroad locomotive engineers, and those schedules are consistent with current Federal regulations. The schedules directly cause backwards rotation of shift start times and progressive sleep deprivation. The backward rotation of shift start times and progressive sleep deprivation, in turn, cause decreased alertness within and across daily runs. Job performance, including safety sensitive tasks such as sounding the train horn at grade crossings, is clearly compromised as alertness decreases across daily runs. As was previously noted, it has been shown in other occupations, that work schedule variables are important determinants of performance ³⁴. The present study provides direct and compelling evidence to suggest that this is also the case for locomotive engineers.

CONCLUSIONS

Current Federal regulations governing Hours of Service for locomotive engineers allow work schedules that have backwards rotating shift start times and do not allow sufficient sleep ³⁵. Locomotive engineers who work under such schedules can accumulate a progressive sleep debt over a period of days. The present study demonstrates that schedules which have these characteristics are easily composed and do, indeed, result in sleep durations which are considerably less than those obtained by the general population. The locomotive engineers in this study, while working on such schedules, reported progressive decreases in subjective alertness across the duration of the study. Moreover, several aspects of job performance, including safety sensitive tasks, degraded during the same time period. This suggests, that current Federal regulations governing Hours of Service have the potential to allow work schedules which degrade the job performance of locomotive engineers and reduce the safety of railroad operations.

REFERENCES

- Åkerstedt, T. (1988). Sleepiness as a consequence of shift work. Sleep, 11, 17-34.
- Bonnet, M. H., and Arand, D. L. (1995). We are chronically sleep deprived. Sleep, 18, 908-911.
- Bradley, J. V. (1968). Distribution-free statistical tests. Englewood Cliffs, NJ: Prentice-Hall.
- Committee on Energy and Commerce (1994). Report to accompany H.R. 4545. Report 103-655.
- Cziesler, C. A., Moore-Ede, M. C., and Coleman, R. M. (1982). Rotating shift work schedules that disrupt sleep are improved by applying circadian principles. Science, 217, 460-463.
- Dement, W. (1995). Managing fatigue in transportation: Promoting safety and productivity. Fatigue Symposium Proceedings. Washington, DC: National Transportation Safety Board and NASA Ames Research Center.
- Gordon, N. P., Cleary, P. D., Parker, C. E., and Czeisler, C. A. (1986). The prevalence and health impact of shiftwork. American Journal of Public Health, 76, 1225-1228.
- Hobson, J. A. (1989). Sleep. New York: Scientific American Library.
- Hours of Service of Railroad Employees. (1993). 49 CFR 228.
- Kuehn, G. I. (1996). Enginemen stress and fatigue: Phase II (Report No. DOT/FRA/ORD-96/05). Washington, DC: Federal Railroad Administration, U. S. Department of Transportation.
- National Transportation Safety Board (1991). Railroad accident report: Atchison, Topeka and Santa Fe Railway Company (ATSF) freight trains ATSF 818 and ATSF 891 on the ATSF Railway, Corona, California, November 7, 1990. (Report Nos. PR91-916303, NTSB/RAR-91/03). Washington, DC: National Transportation Safety Board.
- McFather, L. D. (July 16, 1991). Testimony before The Senate Subcommittee on Surface Transportation, Hearing on Reauthorization of Railroad Safety Programs.
- Monk, T. H. (1989a). Human factors implications of shiftwork. International Reviews of Ergonomics, 2, 111-128.
- Monk, T. H. (1989b). A visual analogue scale technique to measure global vigor and affect. Psychiatry Research, 27, 89-99.

- Monk, T. H., and Folkard, S. (1985). Shiftwork and performance. IN S. Folkard and T. H. Monk (Eds.), Hours of Work - Temporal Factors in Work Scheduling. New York: Wiley. Pp. 239-252.
- Moore-Ede, M. C., Sulzman, F. M., and Fuller, C. A. (1982). The clocks that time us: Physiology of the circadian timing system. Cambridge, MA: Harvard University Press.
- Naitoh, P. (1982). Chronobiologic approach for optimizing human performance. IN F. M. Brown and R. C. Graeber (Eds.), Rhythmic aspects of behavior. Hillsdale, NJ: Erlbaum. Pp. 41-104.
- Pollard, J. K. (1991). Issues in locomotive crew management and scheduling (Report Nos. DOT/FRA/RRP-91-01 and DOT/VNTSC-FRA-91-1). Washington, DC: Federal Railroad Administration, U. S. Department of Transportation.
- Pollard, J. K. (1996). Locomotive engineer's activity diary. (Report Nos. DOT/FRA/RPP-9602, DOT-VNTSC-FRA-96-12). Washington, DC: Federal Railroad Administration, U. S. Department of Transportation.
- Smiley, A. M. (1990). The Hinton train disaster. Accident Analysis & Prevention, *22*, 443-455.
- Smith, M. J., and Colligan, M. J. (1982). Health and safety consequences of shift work in the food processing industry. Ergonomics, *25*, 133-144.
- Smith, M. J., Colligan, M. J., Frock, I. J., and Tasto, D. L. (1979). Occupational injury rates among nurses as a function of shift schedule. Journal of Safety Research, *11*, 181-187.
- Webb, W. B. (1992). Sleep, the gentle tyrant. Anker Publishing.
- United States General Accounting Office. (1992). Railroad safety: Engineer work shift length and schedule variability (Report No. GAO/RCED-92-133). Washington, DC: United States General Accounting Office.

AUTHOR NOTES

Garold R. Thomas is the Project Manager for the Engineman Workload, Stress and Fatigue project within the Office of Research and Development. Mr. Thomas has BSCE and MSCE degrees from West Virginia University. Mr. Thomas is the FRA's representative to the U.S. Department of Transportation's Human Factors Coordinating Committee and the U.S. Department of Transportation's Operator Fatigue Working Group.

Thomas G. Raslear is an Engineering Psychologist in the Office of Research and Development. Dr. Raslear is the technical consultant for psychology, circadian rhythms and statistics for the Engineman Workload, Stress and Fatigue project. Dr. Raslear has a B.S. degree in psychology from the City College of New York and Sc.M. and Ph.D. degrees in psychology from Brown University.

George I. Kuehn is the principal investigator for the Engineman Workload, Stress and Fatigue project at the IIT Research Institute. Dr. Kuehn manages daily operations and the logistics of experimental operations for the RALES simulator at IITRI. Dr. Kuehn has B.S. and M.S. degrees in education from Southern Illinois University and an Ed.D. degree in special education from the University of Northern Colorado.

The research in this report was conducted under Contract No. DTFR53-82-C-00254 at the IIT Research Institute.

The authors acknowledge the support, advice and comments of the following individuals at the Federal Railroad Administration in the preparation of this report: Steven Ditmeyer (Director, Office of Research and Development), Claire Orth (Chief, Equipment and Operating Practices Research Division), and Michael Coplen (Eisenhower Fellow).

Correspondence should be addressed to the authors at the Federal Railroad Administration, Office of Research and Development, RDV -32, 400 Seventh Street, S.W., Washington, DC 20590. E-mail can be addressed to Mr. Thomas at "GAROLD.THOMAS@FRA.DOT.GOV" OR Dr. Raslear at "THOMAS.RASLEAR@FRA.DOT.GOV".

FOOTNOTES

1. The Hours of Service Act of 1907, as amended through 1989.
2. 49 CFR 228.
3. See Pollard, 1991.
4. Ibid.
5. See, for example, Smiley, 1990.
6. United States General Accounting Office, 1992; Pollard, 1996; McFather, 1991, p. 16; National Transportation Safety Board, 1991; Committee on Energy and Commerce, 1994, Section 3.
7. Åkerstedt, 1988; Czeisler, Moore-Ede, and Coleman, 1982; Gordon, Cleary, Parker, and Czeisler, 1986; Monk, 1989a; Monk and Folkard, 1985; Smith and Colligan, 1982; Smith, Colligan, Frockt, and Tasto, 1979.
8. The following table shows the mean ages and years of service of each group of engineers. Group labels are explained in the text.

Age, Years of Service, and Number of Engineers in the Groups.

Group	Age	Years of Service	Number of Engineers
Fast	43.3	14.5	35
Slow	40.5	13.85	20

9. Locomotive engineers must be familiar with the terrain and other features of a territory before they are qualified to operate in that territory without the supervision of a pilot. An experienced and certified locomotive engineer who carries a supervisor of engineers qualification served as pilot during these familiarization runs.
10. This train was selected for two reasons. First, this train represented a very common experience for American locomotive engineers; that is, a unit commodity train of moderate length and

heavy weight. It was, therefore, a type of train which the engineers were likely to have experience with. Second, this train is at once easy to operate and difficult to control. Operation is easy because a loaded unit train has even distribution of car lengths and car weights throughout the train. Slack action forces are more predictable in this type of train than in mixed goods/mixed loads trains. This train is difficult to control because of its great mass relative to its length. Good train handling dictates control of speed by predicting the effects of terrain on the mass of the train. Trains which are heavy can speed up or slow down beyond the control of the engineer if the effects of terrain are not anticipated and corrective control settings made appropriately ahead of time. Vigilance is required in terms of recognizing track locations at which preventive actions are necessary in order to avoid loss of control. The locomotive engineer generally must take these preventative actions prior to any actual change in the state of the train. Once a loaded coal train begins to change its speed as a function of grade, it can be very difficult to have any timely control over that change.

11. The alerter is a device which currently is used in place of the previously used deadman's switch. Unlike the deadman's switch, which required constant operation to prevent an emergency application of the automatic brakes, the modern alerter needs only to be operated periodically. In typical operation, the alerter will flash a light at a period of time which is inversely dependent on the train speed. The alerter is reset by pressing a button on the device, or by changing throttle or brake settings, or by operating the train horn or bell. If none of these actions is performed within a period of time, an audible alarm is sounded. After another period of time, continued failure to act results in an emergency brake application.
12. The actigraph device used was manufactured by Ambulatory Monitoring, Inc. of Ardsley, NY.
13. Action 3 software is produced by Ambulatory Monitoring, Inc. of Ardsley, NY.
14. The Global Vigor and Global Affect scales were adopted from Monk, 1989b.
15. Crew members present in the locomotive cab are supposed to verbally indicate the status of wayside signals as they are approached. This redundancy ensures that signals are properly read, interpreted, and complied with.
16. Kuehn, 1996.
17. Whistle boards are wayside signs which give notice to the engineer that he is approaching a grade crossing. Most railroads have rules which stipulate that horns will be blown at grade crossings.

18. Differences between groups are considered reliable if a statistical test on the data is “statistically significant”. Statistical tests are used to determine if differences between groups are sufficiently large to be considered unlikely (i.e., very low probability) under the hypothesis that the groups are identical in the observed characteristic. “Statistically significant” usually means that a difference between groups (identical in the observed characteristic) is observed which would only occur by chance less than 5% of the time when samples of size N and M are randomly drawn from the groups. If a test yields a statistically significant result, the hypothesis that the groups are identical in the observed characteristic can be rejected in favor of the alternative hypotheses are that the groups are different in the observed characteristic.

For the group comparisons of age and length of service a t-test was used. For age: $t = 1.23, 42 \text{ df}; p > .05$; for length service: $t = 0.27, 44 \text{ df}, p > .05$.

19. Bradley, 1968.

20. For Sleep Duration, the main effect of Group was tested with a Wilcoxon Rank Sum test ($W_n = 176, n = 20, m = 35, p < .01$), the main effect of Runs was tested with a Friedman test ($S = 3022, C = 4, R = 35, \text{df} = 3, \chi_r^2 = 64.76, p < .01$), and the effect of the Group by Run interaction was tested with a Friedman test ($S = 2089, C = 4, R = 35, \text{df} = 3, \chi_r^2 = 44.96, p < .01$).

21. Moore-Ede, Sulzman and Fuller, 1982.

22. For Sleep Onset, the main effect of Group was tested with a Wilcoxon Rank Sum test ($W_n = 289, n = 20, m = 35, p < .01$), the main effect of Runs was tested with a Friedman test ($S = 1750, C = 4, R = 35, \text{df} = 3, \chi_r^2 = 37.5, p < .01$), and the effect of the Group by Run interaction was tested with a Friedman test ($S = 1202, C = 4, R = 35, \text{df} = 3, \chi_r^2 = 25.76, p < .01$).

23. Moore-Ede et al., 1982.

24. The correlation (r) between two measures indicates the extent to which the measures agree. A perfect correlation ($r = 1$) indicates perfect agreement, no agreement would be indicated by a zero correlation ($r = 0$), and a perfect inverse correlation ($r = -1$) indicates that one measure increases as the other decreases. For Global Vigor and alertness, $r = 0.83, n = 80, p < .01$.

25. For Global Vigor on run 3, the Wilcoxon Rank Sum test was used ($W_n = 467, n = 20, m = 34, p > .05$). Predicted alertness was modeled to be identical on run 3. Group effects in Global Vigor across all runs was tested with the Wilcoxon Rank Sum test ($W_n = 485, n = 20, m = 34, p > .05$). The effect Run was tested with the Friedman test ($S = 3834, C = 4, R = 35, \text{df} = 3, \chi_r^2 = 67.66, p < .01$), the effect of the Group by Run interaction was tested with the Friedman test ($S = 2758, C = 4, R = 35, \text{df} = 3, \chi_r^2 = 48.67, p < .01$), and the main effect of

segment was tested with the Friedman test ($S = 1282$, $C = 4$, $R = 10$, $df = 3$, $\chi_r^2 = 34.96$, $p < .01$).

26. Missed horns between groups on run 3 were tested with a (Wilcoxon Rank Sum test ($W_n = 618$, $n = 20$, $m = 35$, $p > .05$). The main effect of Group was tested with the Wilcoxon Rank Sum test ($W_n = 482$, $n = 20$, $m = 35$, $p > .05$), the main effect of Run was tested with the Friedman test ($S = 997$, $C = 4$, $R = 35$, $\chi_r^2 = 11.39$, $p < .01$), and the Group by Run interaction was tested with the Friedman test ($S = 319$, $C = 4$, $R = 35$, $df = 3$, $\chi_r^2 = 3.65$, $p > .05$). The correlation between missed horns and changes in performance predicted by the model was reliable ($r = -0.47$, $n = 80$, $p < .01$).
27. Alerter Response Times between groups on run 3 were tested with a Wilcoxon Rank Sum test ($W_n = 545$, $n = 20$, $m = 35$, $p > .05$). The main effect of Group was tested with a Wilcoxon Rank Sum test ($W_n = 659$, $n = 20$, $m = 35$, $p < .05$), the main effect of Run was tested with the Friedman test ($S = 3561$, $C = 4$, $R = 35$, $df = 3$, $\chi_r^2 = 40.70$, $p < .001$), and the Group by Run interaction was tested with the Friedman test ($S = 3515$, $C = 4$, $R = 35$, $df = 3$, $\chi_r^2 = 40.17$, $p < .001$). The correlation between predicted alertness and the alerter response time was not reliable ($r = -0.19$, $n = 80$, $p > .05$).
28. Fuel Consumption between groups on run 3 was tested with a Wilcoxon Rank Sum test ($W_n = 689$, $n = 20$, $m = 35$, $p < .05$). The main effect of Group was tested with a Wilcoxon Rank Sum test ($W_n = 428$, $n = 20$, $m = 35$, $p < .05$), the main effect of Run was tested with the Friedman test ($S = 629$, $C = 4$, $R = 35$, $df = 3$, $\chi_r^2 = 10.78$, $p < .05$), and the Group by Run interaction was tested with the Friedman test ($S = 265$, $C = 4$, $R = 35$, $df = 3$, $\chi_r^2 = 4.54$, $p > .05$).
29. The correlation between cumulative fuel deviations and predicted alertness was reliable ($r = -0.43$, $n = 60$, $p < .01$).
30. Pollard, 1991; United States General Accounting Office, 1992.
31. Bonnet and Rand, 1995; Hobson, 1989.
32. Naitoh, 1982.
33. The coefficient of determination (r^2), which is the correlation (r) squared, indicates the accuracy of predicting one measure given knowledge of the other measure. In this instance, the correlation between subjective alertness and predicted alertness was 0.83, so that $r^2 = .68$ and the accuracy of predicting subjective alertness from the model is 68%.
34. Åkerstedt, 1988; Czeisler, Moore-Ede, and Coleman, 1982; Gordon, Cleary, Parker, and

Czeisler, 1986; Monk, 1989a; Monk and Folkard, 1985; Smith and Colligan, 1982; Smith, Colligan, Frockt, and Tasto, 1979.

35. In a recent keynote address at a symposium on the management of fatigue in transportation Dr. William Dement (1995) noted that "The average adult sleep requirement is a little over eight hours...". Pollard (1996, p. 11) also indicates that eight hours is the amount of sleep recommended by most experts on the subject. Pollard (1996) presents data collected by Webb (1992) which indicates that the average amount of sleep obtained in the general population is 7.5 hours. Consequently, the amount of sleep obtained by the "slow" and "fast" groups is less than that which is either recommended or generally obtained. Hence, it is not sufficient.

APPENDIX I.

SIMPLE SLEEP DEBT AND CIRCADIAN RHYTHM MODEL

Circadian rhythm

$$A[\cos(\frac{2\pi(T-\phi)}{24})]+100P$$

A is the amplitude, ϕ is the phase shift, and P is the offset.

$$P = \exp -(D/24)$$

D is the sleep debt. If $D=0$, $P=1$, as D increases, $P \rightarrow 0$.

Sleep debt is accumulated while awake. While awake it is the sum of time awake and the value of D remaining after sleep.

Sleep debt is decreased during sleep. Time asleep is subtracted from the sum of the (hours awake divided by 4) and (cumulative sleep debt prior to sleep divided by 4).

For the present application of the model, the following parameter values were used:

$$A = 5$$

T is time of day in 24 hour notation

$$\phi = 12$$