FHWA report number: FHWA-MC-97-001
TC report number: TP 12876E

COMMERCIAL MOTOR VEHICLE DRIVER FATIGUE AND ALERTNESS STUDY

TECHNICAL SUMMARY

Prepared for

Federal Highway Administration
U.S. Department of Transportation

Trucking Research Institute
American Trucking Associations Foundation

Transportation Development Centre
Safety and Security
Transport Canada

November 1996

This report is also available in French from Transport Canada under the title “Étude sur la fatigue et la vigilance chez les conducteurs de véhicules utilitaires - Sommaire technique” TP 12876F.

Ce rapport est aussi disponible en français à Transports Canada sous le titre «Étude sur la fatigue et la vigilance chez les conducteurs de véhicules utilitaires - Sommaire technique» TP 12876F.

Technical Report Documentation Page

1. Report No.
FHWA-MC-97-001

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle
Commercial Motor Vehicle Driver Fatigue and Alertness Study:
Technical Summary

5. Report Date
November 1996

6. Performing Organization Code

7. Author(s)
C.D. Wylie, T. Shultz, J.C. Miller, M.M. Mitler, R.R. Mackie

8. Performing Organization Report No.

9. Performing Organization Name and Address
Essex Corporation
5775 Dawson Avenue
Goleta, California 93117

10. Work Unit No. (TRAIS)

11. Contract or Grant No.
DTFH61-89-C-096, DTFH61-90-C-053

12. Sponsoring Agency Name and Address
Federal Highway Administration
400 7th Street, SW
Washington, DC 20590

13. Type of Report and Period Covered
Technical Summary

14. Sponsoring Agency Code

15. Supplementary Notes
D.M. Freund -- Project Manager and Contracting Officer's Technical Representative
Project Coordinators: S. Vespa (Transport Canada) and C. E. Woodle (Trucking Research Institute)
This report is being published simultaneously by Transport Canada, Publication No. TP 12876E.

16. Abstract

This is the technical summary of the research report, Commercial Motor Vehicle Driver Fatigue and Alertness Study by Wylie et. al., the largest and most comprehensive over-the-road study on this subject ever conducted in North America.

The data collection involved eighty drivers in the U.S. and Canada who were monitored over a period of sixteen weeks. A number of work-related factors thought to influence the development of fatigue, loss of alertness, and degraded driving performance in commercial motor vehicle drivers was studied within an operational setting of real-life, revenue-generating trips. These included: the amount of time spent driving during a work period; the number of consecutive days of driving; the time of day when driving took place; and schedule regularity.

In Section 1 of the Technical Summary, the reader is provided with some extracts from the technical literature on the involvement of fatigue in crashes, a historical summary of the U.S. Department of Transportation's focus on commercial motor vehicle driver fatigue and the background to this study. The study's overall objectives and the approach used in their attainment are also provided. Section 2 presents the conclusions drawn from the literature review conducted in preparation for this study and considered in the formulation of the study's own conclusions and recommendations. Section 3 presents the study's methodology and data collection methods, while Section 4 presents the study's results, conclusions, and recommendations.

For the amount of sleep and the four to five days of driving observed for each driver in the study, it was found that the strongest and most consistent factor influencing driver fatigue and alertness was time of day; drowsiness, as observed in video recordings of the driver's face, was markedly greater during night driving than during daytime driving. The number of hours of driving (time-on-task) and cumulative number of days were not strong or consistent predictors of observed fatigue. Numerous other findings are provided relating to scientific methodologies and fatigue

17. Key Words
Truck driver, fatigue, alertness, CMV drivers, vigilence, shiftwork, driver

18. Distribution Statement
No restrictions. Available through the National Technical Information Service, Springfield, VA 22151

19. Security Classif. (of this report)
Unclassified

20. Security Classif. (of this page)
Unclassified

21. No of Pages

22. Price

Form DOT F 1700.7 (8-72)
(PF V2.1, 6/30/92)

Reproduction of completed page authorized

Transport Canada Transports Canada

PUBLICATION DATA FORM

1. Transport Canada Publication No.
TP 12876E

2. Project No.
8274, 8277

3. Recipient's Catalogue No.

4. Title and Subtitle
Commercial Motor Vehicle Driver
Fatigue and Alertness Study:
Technical Summary

5. Publication Date
November 1996

6. Performing Organization Document No.

7. Author(s)
C.D. Wylie, T. Shultz, J.C. Miller, M.M. Mitler, and R.R. Mackie

8. Transport Canada File No.
ZCD1450-110-9

9. Performing Organization Name and Address
Essex Corporation
5775 Dawson Avenue
Goleta, CA 93117

10. PWGSC File No.

11. PWGSC or Transport Canada Contract No.
T8200-3-3531/01-XSD

12. Sponsoring Agency Name and Address
Transportation Development Centre (TDC)
800 René Lévesque Blvd. West
6th Floor
Montreal, Quebec
H3B 1X9

13. Type of Publication and Period Covered
Technical Summary

14. Project Officer
S. Vespa

15. Supplementary Notes (Funding programs, titles of related publications, etc.)

Research was conducted under a U.S. Department of Transportation and Transport Canada research program agreement, with FHWA as the contracting agency and D.M. Freund as FHWA's Project Manager and Contracting Officer's Technical Representative. Financial contribution of Trucking Research Institute (TRI) of the American Trucking Associations Foundation as well as participation on the project Steering Committee of TRI and the Canadian Trucking Research Institute of the Canadian Trucking Association are gratefully acknowledged. This report is simultaneously published by FHWA, Publication No. FHWA-MC-97-001.

16. Abstract

This is the Technical Summary of the research report *Commercial Motor Vehicle Driver Fatigue and Alertness Study* by Wylie et al., dated October 1996, concerning the largest and most comprehensive over-the-road study on this subject ever conducted in North America.

The data collection involved eighty drivers in the U.S. and Canada who were monitored over a period of sixteen weeks. A number of work-related factors thought to influence the development of fatigue, loss of alertness and degraded performance in CMV drivers was studied within an operational setting of real-life, revenue-generating trips. These included: the amount of time spent driving during a work period; the number of consecutive days of driving; the time of day when driving took place; and schedule regularity.

In Section 1 of the Technical Summary, the reader is provided with some extracts from the technical literature on the involvement of fatigue in crashes, a historical summary of the U.S. Department of Transportation's focus on commercial motor vehicle driver fatigue and the background to the study. The study's overall objectives and the approach used in their attainment are also provided. Section 2 presents the conclusions drawn from the literature review conducted in preparation for the study and considered in the formulation of the study's own conclusions and recommendations. Section 3 presents the study methodology and data collections methods, while Section 4 presents the study's results, conclusions and recommendations.

For the amount of sleep and the four to five days of driving observed for each driver in this study, it was found that the strongest and most consistent factor influencing driver fatigue and alertness was time-of-day; drowsiness, as observed in video recordings of the driver's face, was markedly greater during night driving than during daytime driving. The number of hours of driving (time-on-task) and cummulative number of days were not strong or consistent predictors of observed fatigue. Numerous other findings are provided relating to scientific methodologies and fatigue countermeasure concepts.

17. Key Words
Truck driver, fatigue, alertness, CMV drivers,
vigilance, shift work, driver

18. Distribution Statement
Limited number of copies available from the
Transportation Development Centre

19. Security Classification (of this publication)
Unclassified

20. Security Classification (of this page)
Unclassified

21. Declassification (date)
—

22. No. of Pages
x, 59

23. Price
—

CDT/TDC 79-005
Rev. 96

Canada

Transports Canada Transport Canada

FORMULE DE DONNÉES POUR PUBLICATION

1. N° de la publication de Transports Canada

TP 12876E

2. N° de l'étude

8274, 8277

3. N° de catalogue du destinataire

4. Titre et sous-titre

Commercial Motor Vehicle Driver
Fatigue and Alertness Study:
Technical Summary

5. Date de la publication

Novembre 1996

6. N° de document de l'organisme exécutant

7. Auteur(s)

C.D. Wylie, T. Shultz, J.C. Miller, M.M. Mitler, et R.R. Mackie

8. N° de dossier - Transports Canada

ZCD1450-110-9

9. Nom et adresse de l'organisme exécutant

Essex Corporation
5775 Dawson Avenue
Goleta, CA 93117

10. N° de dossier - TPSGC

11. N° de contrat - TPSGC ou Transports Canada

T8200-3-3531/01-XSD

12. Nom et adresse de l'organisme parrain

Centre de développement des transports (CDT)
800, boul. René-Lévesque Ouest
6e étage
Montréal (Québec)
H3B 1X9

13. Genre de publication et période visée

Sommaire technique

14. Agent de projet

S. Vespa

15. Remarques additionnelles (programmes de financement, titres de publications connexes, etc.)

Recherche menée sous les auspices d'un accord de programme de recherche entre le Department of Transportation des États-Unis et Transports Canada, avec la FHWA agissant comme organisme contractant et D.M. Freund à titre de gestionnaire de projet de la FHWA et représentant technique du bureau du contractant. L'apport financier fourni par le Trucking Research Institute (TRI) de l'American Trucking Associations Foundation de même que la participation, au Comité de direction de ce projet, du TRI et de l'Institut canadien de recherche sur le camionnage de l'Association canadienne du camionnage méritent bien d'être soulignés. Ce rapport est publié en même temps par la FHWA, Publication No. FHWA-MC-97-001.

16. Résumé

Sommaire technique du rapport de recherche intitulé *Étude sur la fatigue et la vigilance chez les conducteurs de véhicules utilitaires*, par Wylie et coll., daté d'octobre 1996, la plus vaste et la plus complète étude sur ce sujet à être menée en Amérique du Nord.

Quatre-vingt conducteurs aux États-Unis et au Canada, observés sur une période de 16 semaines, ont été mis à contribution pour la collecte de données. Certains facteurs liés au travail et perçus comme pouvant avoir une influence sur l'apparition de la fatigue, la perte de vigilance et la dégradation de la performance chez les conducteurs de véhicules utilitaires ont été étudiés dans le cadre d'expérimentations menés en service payant. Parmi ces facteurs : le nombre d'heures consacrées à la conduite durant une période de travail, le nombre de jours consécutifs au volant, la période du jour au volant et la régularité des horaires.

La partie 1 du Sommaire technique, qui trace les grandes lignes de l'étude, reprend certains extraits de la documentation technique produite sur le rapport entre la fatigue et l'occurrence de collisions; fait une synthèse de l'intérêt continu qu'a porté le Department of Transportation des États-Unis au problème de la fatigue chez les conducteurs de véhicules utilitaires; place l'étude dans son contexte historique et énonce les objectifs globaux ainsi que la méthode utilisée pour les atteindre. La partie 2 porte sur les conclusions tirées de la revue approfondie de la documentation ayant servi à la préparation de l'étude et à la formulation de ses propres conclusions et recommandations. La partie 3 présente la méthode choisie de même que les modalités de collecte de données adoptées pour l'étude, tandis que la partie 4 porte plus précisément sur les résultats, les conclusions et les recommandations.

Pour ce qui est de la quantité de sommeil et les quatre ou cinq jours de conduite observés chez chaque conducteur dans cette étude, il a été noté que le facteur le plus dominant et le plus constant influençant la fatigue et la perte de vigilance chez les conducteurs était la période du jour. La somnolence telle qu'observée dans les enregistrements vidéo des expressions faciales était considérablement plus marquée lors de la conduite de nuit que durant la conduite de jour. Le nombre d'heures de conduite (heures au volant) et le nombre cumulatif de jours n'étaient pas des prédicteurs dominants ou constants de la fatigue observée. De nombreuses autres découvertes figurent également dans ce document, notamment les méthodologies scientifiques et les concepts de contre-mesures à la fatigue.

17. Mots clés

Camionneur, fatigue, vigilance, conducteurs de véhicules utilitaires, travail par roulement, conducteur

18. Diffusion

Le Centre de développement des transports dispose d'un nombre limité d'exemplaires.

19. Classification de sécurité (de cette publication)

Non classifiée

20. Classification de sécurité (de cette page)

Non classifiée

21. Déclassification (date)

—

22. Nombre de pages

x, 59

23. Prix

—

Canada

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

COMMERCIAL MOTOR VEHICLE
DRIVER FATIGUE AND ALERTNESS STUDY

TECHNICAL SUMMARY

This Technical Summary presents, in four sections, the research conducted under the Federal Highway Administration's (FHWA's) Commercial Motor Vehicle Driver Fatigue and Alertness Study (DFAS). In the introduction to the study (Section 1), the reader is provided with some extracts from the technical literature on the involvement of fatigue in crashes, a historical summary of the U.S. Department of Transportation's focus on commercial motor vehicle (CMV) driver fatigue and the background to this study, as well as the study's overall objectives and the approach used in their attainment. Section 2 presents the conclusions drawn from the extensive literature review conducted in preparation for this study and considered in the formulation of the study's own conclusions and recommendations. Section 3 presents the study methodology and data collection methods, while Section 4 presents the study's results, conclusions and recommendations.

SECTION 1. INTRODUCTION

FATIGUE AND CRASHES

A 1965 press release from the office of the Oklahoma Turnpike Authority (cited by Case and Hulbert, 1970; Harris and Mackie, 1972)[*] stated that 22% of the 2,128 motor vehicle accidents occurring between 1953 and 1964 were the result of the driver being asleep at the wheel. These accidents reportedly accounted for 48% of the traffic fatalities. Since that time, there have been several other studies addressing driver fatigue and loss of alertness in motor vehicle accidents.

The National Transportation Safety Board (1995) examined single-vehicle truck accidents that were likely to have been fatigue-related to determine the role of specific factors, such as driver work-rest cycles, in fatigue-related accidents. The selection was further limited to those accidents in which the driver survived and in which the previous 96 hours could be reconstructed. Of the 107 accidents selected, 58% were considered to have fatigue as a probable cause, while the remainder were considered not fatigue-related. Seventy of the 107 accidents occurred between 2200 and 0800 and, of these, about 74% (52 of 70) were considered fatigue-related. Of the accidents that occurred between 0800 and 2200, 73% (27 of 37) were not considered fatigue-related. Nineteen of the 107 drivers stated that they fell asleep while driving.

---

[*] The list of references is given in the main report (FHWA report no. FHWA-MC-97-001; Transport Canada report no. TP12875E)

In assessing findings from case studies involving a limited number of specifically selected accident cases, care must be taken not to generalize their results and project them to the total population of accidents. Haworth et al. (1988) pointed out the difficulty of determining the strength of the relationship between fatigue and crashes because interpretation of statistics requires judgments be made as to control of exposure for crash risk, how separate classes of vehicles and crashes are treated, how data were collected, and other factors (p. 71). However, they also noted that fatigue-related crashes have the potential to be more severe than other crashes and are often fatal.

In a study of heavy vehicle drivers in Australia, Haworth, Heffernan and Horne (1989) estimated that fatigue was a contributing factor (whether on the part of a car driver or a truck driver) to between 9 and 20% of fatal accidents involving trucks. In a review of in-depth studies, police reports and driver surveys from around the world, Haworth, Triggs and Grey (1988) concluded that the contribution of fatigue to accidents in articulated vehicles probably ranges between 5 and 10% of all crashes, about 20-30% of casualty crashes and about 25-35% of fatal crashes. The contribution of fatigue may even reach 40-50% in particular types of crashes, for example fatal single-vehicle semi-trailer crashes. MacDonald (1984) notes that fatigue and falling asleep are particularly associated with single-vehicle accidents and rear-end collisions, and these types of accidents are disproportionately common at night. Single-vehicle accidents have higher death rates for truck drivers than accidents involving other types of vehicles.

Knipling and Wang (1994) examined the incidence of crashes related to driver drowsiness/fatigue. They stated that National Highway Traffic Safety Administration (NHTSA) General Estimates System (GES) statistics for 1989-93 showed about 1% of police-reported crashes (including both cars and trucks) cited driver drowsiness/fatigue. They suggested this may be an underestimate for a number of reasons: lack of a specific checkoff box for fatigue on some state police report forms, lack of firm evidence on which to base a police finding of fatigue, lack of awareness on the part of the driver of his/her own fatigue, and the significant number of crashes involving "drift out of lane" which are not cited as drowsiness-related. Knipling and Wang indicated that data from the 1989-93 Fatal Accident Reporting System (FARS) show that 3.6% of fatal crashes were cited as involving drowsiness/fatigue, but again stated that this may be an underestimate for many of the same reasons. Knipling and Wang also cite Deering's 1994 review of 1,000 Michigan Police Accident Reports, which found 1% attributed to "dozing," 17% more attributed to "daydreaming"/distraction, and another 18% to "looked but didn't see." They cited Treat, Tumbas, McDonald et al. (1979) as finding that 56% of crashes involved "recognition errors" as a certain or probable factor, and note that the role of fatigue in such recognition failures and other mental errors resulting in crashes is currently unknown.

The problem of driver fatigue is clearly established in the case of drivers of commercial vehicles whose responsibilities often involve irregular schedules and many driving hours, but the conditions that lead to fatigue being a cause of traffic accidents certainly apply to automobile drivers as well. To support the work of the New York State Task Force on the Impact of Fatigue on Driving, a telephone survey of 1,000 randomly selected drivers, representative of the New York State population in age, gender and county of residence, was conducted (McCartt, Pack and Riborer, 1995). Of the survey drivers, 55% said that they had driven drowsy in the last year; 2.5% said that they frequently drive while drowsy; 23% had fallen asleep at the wheel without crashing; 1.9% report a crash due to drowsiness; and 2.8% had a crash due to falling asleep at the wheel.

It has been reported that 31% of drivers who had experienced drowsiness were initially unaware of the onset of their drowsiness (Skipper and Wierwille, 1986). Dingus, Hardy and Wierwille (1987) concluded that two of the leading causes of automobile accidents are driver impairment due to alcohol and drowsiness. A relatively large percentage of these accidents is believed to occur because drivers are unaware of the degree to which they are impaired.

Despite the variation in quantitative estimates of the role of driver fatigue in highway safety, there is little doubt that fatigue is an important issue. Although driver drowsiness/fatigue is cited on police accident reports as a causal factor in a relatively small percentage of truck accidents, it is believed to play a larger role than cited due to underreporting and to subtle effects on driver performance. Researchers suggest that the contribution of fatigue is likely to be underestimated in that other accident causes, such as inattention, distraction/daydreaming, looked but didn't see, are cited instead of fatigue although these may, in some cases, arise from a fatigued state. Case studies suggest that fatigue plays a more significant role when truck fatal and injury accidents are considered.

HISTORY OF DOT'S FATIGUE FOCUS

The maximum amount of time that commercial motor vehicle (CMV) drivers operating in interstate commerce may drive their vehicles is specified in Title 49, Code of Federal Regulations, Part 395. In Canada, it is limited under the federal "Commercial Vehicle Drivers Hours of Service Regulations, 1994"; SOR/DORS/94-716, 15 November 1994. The U.S. regulations were originally developed in 1935 by the Interstate Commerce Commission (ICC). On April 25, 1938, the ICC requested the United States Public Health Service (USPHS) to conduct an investigation into the hours of service (HOS) of drivers of commercial motor vehicles operating in interstate commerce. This was the first scientific study to address fatigue relating to hours of service. The USPHS found that "it would...appear that a reasonable limitation of the HOS would, at the very least, reduce the number of drivers on the road with very low functional efficiency. This, it might reasonably be inferred, would act in the interest of highway safety." (Jones et al., 1941) No further study was undertaken by USPHS or the ICC. In December, 1967, the ICC's responsibilities concerning CMV driver and vehicle safety were transferred to the then Bureau of Motor Carrier Safety (now Office of Motor Carriers) of the Federal Highway Administration (FHWA), an agency within the then newly-created U.S. Department of Transportation (DOT).

The DOT has devoted considerable resources to addressing this issue. In the 1970's, three major field research studies were conducted to assess the influences on driver alertness of driving time (Mackie and Miller, 1978; Harris and Mackie, 1972), heat, noise, and vibration (Mackie et al., 1974), and physical effort expended in loading and unloading cargo (Mackie and Miller, 1978).

Driver alertness was measured using a broad-spectrum approach incorporating driving task performance patterns and the drivers' physiological and behavioral responses as well as self-evaluations. Patterns of accidents were also studied to determine whether there was a relationship between driving time and accident occurrence (Mackie and Miller, 1978; Harris, 1977; Harris and Mackie, 1972). Although causal relationships were noted, they were not considered strong enough to justify changes proposed by the DOT to the hours-of-service regulations in 1979. Driver fatigue became a safety focus again in the mid-1980s. A 1987 Office of Technology Assessment report, "Gearing Up for Safety," pointed to driver fatigue as a growing highway safety concern. Under the Truck and Bus Safety and Regulatory Reform Act of 1988, the Congress directed the DOT to conduct research to determine the relationship, if any, among Federal hours-of-service regulations for operating commercial vehicles, operator fatigue, and the frequency of serious accidents involving CMVs.

In November, 1988, the FHWA held a Symposium on Truck and Bus Driver Fatigue to discuss what was known about fatigue and fatigue-related accidents and to propose research on that subject. The conference brought together experts from the motor carrier industry, the scientific and medical communities, law enforcement, and public policy. The Driver Fatigue and Alertness Study was a direct result of the Symposium recommendations.

Driver fatigue continues to be a concern. This was made clear at the National Truck and Bus Safety Summit in March 1995 sponsored by the FHWA to identify the major safety issues facing the motor carrier industry. Representatives of the many facets of the motor carrier industry and highway safety community came to a formally-structured consensus that fatigue is the highest-priority issue affecting the safety of motor carriers, and that "drivers, dispatchers, trucking company management, and OMC [FHWA's Office of Motor Carriers] need more factual information about fatigue, and how factors under their control affect fatigue impairment risks" (FHWA, 1995). The present study is an important step toward this goal.

BACKGROUND TO THIS STUDY

The hours-of-service regulations for CMV drivers are among the most important factors that affect the productivity of the trucking industry and its resulting impacts on the cost of goods transported by truck to consumers. They are also of utmost importance to the safety of trucking operations -- both for the safety of CMV drivers themselves and of others sharing the roads with them.

At the time this study was initiated in 1989, it was considered essential by government, industry and academia to undertake an exploratory research program that had high face validity (e.g., operationally valid and statistically reliable) in order to advance our understanding, with a high degree of confidence, of the important factors influencing CMV driver fatigue and loss of alertness, as well as their relative importance.

Research funding for this study was made available by DOT subsequent to the 1988 Symposium held by the FHWA. A competitive contract was awarded to Essex Corporation, Columbia, Maryland, in September, 1989.

In September, 1990, a second, companion contract for collection and analysis of an expanded array of physiological data from the same driver group was awarded to the Trucking Research Institute of the ATA Foundation. The subcontracts for this research were awarded in January, 1992. Essex was made responsible for organizing the physiological data collection and integrating the results with the driving performance data; Scripps Clinic and Research Foundation, La Jolla, California, and Miller Ergonomics, Imperial Beach, California, were responsible for reducing and interpreting the physiological data and providing related technical advice to Essex.

In the summer of 1993, a research agreement between the DOT and Transport Canada was expanded to include sharing of resources and costs for the conduct of field data collection activities in Canada. Transport Canada had been participating in the consultations undertaken by the FHWA since the fall of 1989.

Motor carriers were recruited during 1991 and 1992, and the field data collection work was conducted in both Canada and the U.S. in 1993. During the field data collection program, the sleep laboratory work and associated subject preparation were conducted by the Sleep Disorders Clinic of the Deaconess Hospital, St. Louis, Missouri, for the U.S. operations and the Sleep Disorders Centre of Metropolitan Toronto, Ontario, for the Canadian operations.

STUDY OBJECTIVES

The primary goals of this research were to investigate the effects on safety-related driving performance of the primary factors commonly thought to lead to the development of fatigue and loss of alertness of commercial vehicle drivers; to investigate the relative importance of their effects; to establish objective and measurable relationships between the primary fatigue-producing factors and driving performance; and to identify effective and efficient countermeasures based on the study's findings.

Secondary goals of the research were to investigate the potential for developing a driver alertness monitoring system based on vehicle and driver based measurements; to identify an effective reduced data set for more economically conducting future fatigue research based on collected field data; and, to provide a data set that could be used for validating future fatigue research based on driving simulators.

The results of this study were further intended to provide a scientifically valid basis to determine the potential for revisiting the current hours-of-service requirements, which have been essentially unchanged for more than fifty years.

OVERVIEW OF STUDY APPROACH

The study was designed to investigate, in an operational context, the work-related factors that were considered to lead to the development of fatigue and loss of alertness in commercial motor vehicle drivers and that affected driving safety. These factors were: the amount of time spent driving on a continuing basis in a duty period (e.g., relating to acute fatigue); the number of consecutive work-rest cycles (e.g., relating to cumulative fatigue); the time of day that driving took place (e.g., relating to circadian fatigue); the number of hours slept in principal sleep periods (e.g., relating to sleep deficits); and schedule regularity (e.g., relating to circadian effects on driving performance and on the amount of sleep obtained and its quality). Not only were these work-related factors important from the safety perspective, they were also considered critical to the productivity of this transport mode.

Because of the difficulty of obtaining results on an absolute basis, the study was designed to examine relative impacts of fatigue factors. Budget and time constraints as well as difficulties associated with recruiting fleets having the required characteristics limited the number of factors that could be examined in a rigorous manner. The study design also had to accommodate Canadian drivers operating in their home jurisdiction in order to take advantage of the different Canadian hours of service rules. These rules made the desired comparisons possible using company drivers working in a familiar operational environment. The data collection program was thus established to address the work-related factors leading to fatigue and loss of alertness as follows.

Amount of time spent driving per trip: The study observed trips of both 10 and 13-hour duration in order to permit within and between Condition comparisons at the respective legal limits in the U.S. and Canada.

Number of consecutive days spent driving: Four days and five days of consecutive driving were incorporated to permit within and between Condition comparisons at the legal limits (Canada and the U.S. have similar weekly caps on driving time but different limits on duration of driving time, which permitted four and five days of operation in Canada and the U.S., respectively).

Time of day that driving takes place: Because of the long driving times being studied, there are confounding effects between driving duration and the time of day at which driving takes place. Since these are important effects to assess, the study design had to deal with the confounding effects in a way that permitted valid comparisons. This was done by incorporating two driving schedules into the data collection that were matched in duration of drive but that contrasted the time of day that work was performed. One driving schedule involved starting the drive at about midnight, which is near the nominal circadian low point of 4 a.m., and ending in the early afternoon, which is near the nominal circadian high point of 4 p.m., after having accumulated 13 hours of driving time. The second schedule involved starting the drive at about noon, which is near the nominal circadian high point, and ending in the early morning, which is near the circadian low point, after having accumulated 13 hours of driving time. This approach permitted discrimination between the relative effects arising from duration of drive and time of drive. A further comparison could then also be made with a baseline or reference schedule involving 10 hours of driving.

Schedule regularity: This factor was not extensively studied since it would have required substantially more time and resources than were available. As an exploratory investigation, one schedule was incorporated which had, with each consecutive day, a 3-hour backward rotating start time, from an early morning start on the first day. This was considered to be representative of a large number of long-distance, time-sensitive operations in the U.S. All other schedules had regular start and end times in order to minimize the confounding effects of schedule irregularity.

Amount of sleep obtained: This factor was not controlled as a study variable. In the operational environment of this study, sleep limitations were perceived to have ethical and safety implications. The intent was to examine driving performance in a quasi-naturalistic setting: the drivers would obtain their sleep in a hospital or motel room equipped with a standard bed and instruments to provide a controlled data-collection environment. The study protocol allowed the drivers to determine their own bedtimes, although they were advised of the need to get adequate sleep. The study design was developed to provide drivers with off-duty time as required by their respective U.S. and Canadian regulations. It was thought that this approach would also provide an understanding of driver sleep patterns in different shifts, in an operational setting. Although this approach risked confounding study results because of the potentially different driver sleep times that would result, it was thought necessary to respond to industry's strong desire to have the study as representative as possible of real world conditions. It was also thought that, even if the amount of sleep were to be controlled, the different schedules would result in circadian impacts on the quality of sleep obtained. Drivers would manage their own level of sleep and, indirectly, their performance level. However, drivers would be monitored continuously during sleep in order to relate sleep quantity and quality to their driving performance, test performance, and self reports.

STUDY PARTNERSHIPS

The FHWA held several technical consultation meetings during the planning phase to solicit viewpoints on perceived relationships among the scientific goals of the study, the declared research needs of various stakeholders, the proposed experimental protocols, and the opportunities to apply those protocols toward collecting information during "real world" motor carrier operations -- revenue-producing runs. The meetings brought together representatives from the trucking industry, drivers, law enforcement officials, scientists, and policy experts.

The study was a venture developed and conducted with the help of the motor carrier industry. Industry participated by providing input during the project's planning phase and also provided the motor carriers and drivers who were subjects of the field data collection. The study was also a public-private partnership, and an international one. The Trucking Research Institute of the American Trucking Associations (ATA) Foundation and Canada's department of transport, Transport Canada, funded a significant portion of the data collection and analysis effort as well as preparation of the final project report. The ATA, the National Private Truck Council, the International Brotherhood of Teamsters, and the Owner-Operator Independent Drivers Association provided considerable input in the public meetings. These organizations, as well as the Canadian Trucking Association and the Private Motor Truck Council of Canada, paved the way for recruiting motor carriers and drivers and provided ongoing technical and operational support to the research effort.

SECTION 2. CONCLUSIONS DRAWN FROM LITERATURE REVIEW

The study final report presents an extensive and detailed chapter on the results of the literature review conducted as part of this study. For reasons of brevity, only the conclusions are presented in the following:

1. While driver drowsiness/fatigue is cited on police accident reports as a causal factor in a relatively small percentage of truck accidents, it is believed to play a larger role than cited due to underreporting and to subtle effects on driver performance. Case studies suggest that it plays a significant role in accidents that result in fatalities and injuries.

2. Although there is some disagreement about the definition of driver fatigue, most investigators agree that a reasonable operational definition includes time-correlated deterioration in driving performance, physiological state of arousal, and subjective feelings of sleepiness or tiredness.

3. The consequences of driver fatigue are believed to include:
    Increased lapses of attention
    Increased information processing and decision making time
    Increased reaction time to critical events
    More variable and less effective control responses
    Decreased motivation to sustain performance
    Decreased psychophysiological arousal (e.g., brain waves, heart action)
    Increased subjective feelings of drowsiness or fatigue
    Decreased vigilance (e.g., watchfulness)
    Decreased alertness (e.g., readiness)

4. The primary causes of driver fatigue are long periods of driving, circadian low points and sleep debt. Fatigue effects have also been associated with:
    Rotating schedules
    Two-person or team "sleeper" operations
    Monotonous driving environments

Driving in darkness
Adverse weather conditions
Alcohol and drugs
Physical work, in addition to driving
Noise, vibration and heat

5. Research on driver fatigue typically employs several measures of driver performance and several simultaneous recordings of the driver's state of alertness or physiological arousal. In some studies, the driver's self-evaluations are used to augment the objective measures. Driver performance is measured in connection with both the primary driving task (vehicle control) and, frequently, on a variety of secondary tasks as well. The purpose of the latter is to provide a convenient, objective means of detecting changes in the driver's reaction time to critical signals that may reflect fatigue effects. A great variety of secondary tasks have been used, with mixed success. Some may interfere with the primary control task.

6. There have been two main thrusts of driver fatigue research: (a) studies aimed at measuring the extent and time-course of loss of driver alertness (primarily with the objective of evaluating hours of service regulations) and (b) studies aimed at developing real-time, in-vehicle drowsy driver detection and alerting systems. Both types of studies have generally found the same driver performance and physiological status variables to be useful indicators of driver fatigue.

7. Much recent research on driver fatigue has been performed in driving simulators. While this permits a degree of experimental control not possible in over-the-road studies, many variables that affect driver fatigue may operate differently in the real world than in simulators (if they are represented in simulators at all). Therefore, significant findings about driver fatigue from simulator studies require real-world validation.

8. Many countermeasures aimed at minimizing the problem of fatigue-related accidents have been proposed, although few have been the subject of operational testing. Broadly, countermeasures fall into the following categories:

   Alarm systems, based on changes in driver performance, or level of psychophysiological arousal, or both;

   Alertness maintainers, in the form of driving hour regulations, obligatory rest stops, napping, training about factors that cause driver fatigue, and certain devices installed in the driving environment.

9. For countermeasures to be effective they must meet a number of operational and driver acceptance criteria presented in this review. No existing device or procedure has, to date, been formally assessed against these criteria. Informal analysis suggests that driver alertness monitoring devices developed thus far would fail to meet many of them. On the brighter side, progress is being made toward in-vehicle driver alertness monitoring devices that meet at least two of the most important criteria: sensitivity to driver impairment and acceptable false alarm rates (Wierwille et al., 1994). However, much work remains to be done.

SECTION 3. METHODS

This section begins with an overview of the data collection approach and moves on to describe the study design and field operations, including schedule and driver characteristics. The study's data collection methodology is then described including the driving performance, surrogate test performance, and physiological measures used. Finally, the instrumentation and field procedures are summarized, ending with a description of the collected database.

OVERVIEW

For purposes of this study, we considered the driver's state of fatigue as resulting from working for extended periods, under adverse schedules (including driving and sleeping at times not synchronized with natural body rhythms), and/or from not getting adequate sleep. The goal of the on-the-road data collection was to provide data with which to establish quantitative relationships between the development of driver fatigue and decreasing proficiency (if any) in driving-related tasks.

The independent variables of primary interest were (1) the amount of time spent driving, (2) the time of day that driving was initiated, and (3) the amount of sleep obtained. Driving duration was varied to assess the relative influence of time on task. The time of day that driving was initiated was varied in a planned fashion in order to assess the influence of schedule regularity and circadian effects. Sleep duration was monitored to permit interpretation of driving time and circadian rhythm effects but was not controlled in order to observe drivers’ off-duty rest patterns. The study design was developed to comply with existing U.S. and Canadian hours-of-service regulations.

All data were collected during revenue-producing runs. These runs were the same as the participating drivers normally drove, or could be expected to drive, from their home terminal. The routes included open-road driving through flat or rolling terrain on limited-access highways, and urban driving in the vicinity of their motor carriers' terminals. The highway environments between the 10-hour conditions, in the U.S., and the 13-hour conditions, run in Canada, were matched as closely as possible to minimize differential effects of roadway geometrics, terrain, and traffic volume and mix. While some driving took place during heavy rainstorms, none took place during snowy or icy conditions.

The drivers all drove late-model conventional (engine ahead of cab) tractors in their respective fleets. None of the drivers was required as part of their regular duties to load or unload freight, nor were they required to do so as part of this study.

A broad spectrum of measures is necessary to obtain an adequate picture of the evolution of driver fatigue and driving performance since there is no one "best" measure available, nor is any single measure or class of measures sufficient. In this study we used a variety of direct driving performance measures as well as indirect (surrogate) performance measures, physiological measures, assessments based on visual observation of driver behavior, and driver self-evaluations.

The ideal outcome from the employment of multiple measures is convergent validity; i.e., results from different measures which corroborate each other. Achieving such consistency of results helps to demonstrate both the validity of the individual measures employed as well as general conclusions about the underlying theoretical construct of interest -- driver fatigue. However, because there is no "best" measure of alertness/fatigue and because the different measures employed relate to fundamentally different manifestations of the underlying process (e.g., brain cortex electrophysiology, subjective experience, motor performance, observed facial expression and eye activity), high convergent validity is often difficult to achieve in studies of alertness/fatigue. This is particularly true in a field study setting where uncontrolled variables may to some extent confound measurement protocols. The study methodology included assessment of the consistency of results from different measures. Areas of consistency and inconsistency across different measures are noted as part of the study results and related discussion for major fatigue issues.

The overall data collection program is presented in schematic form in Figure 1 to facilitate the reader's understanding of its various elements and their interrelationships.

STUDY DESIGN AND FIELD OPERATIONS

From the outset, we intended to focus the study on the most demanding operations permissible within current hours-of-service rules, to shed light on driver performance near the limits permitted by those rules. Driving schedules were organized to take advantage of the Canadian 13-hour driving limit, to contrast driver performance under the U.S. 10-hour limit with that under the 13-hour limit. Table 1 summarizes key aspects of the U.S. and Canadian hours-of-service regulations relevant to this study. It was also desirable to evaluate the effects of irregular schedules and night driving, because these are generally thought to lead to fatigue and loss of alertness symptoms.

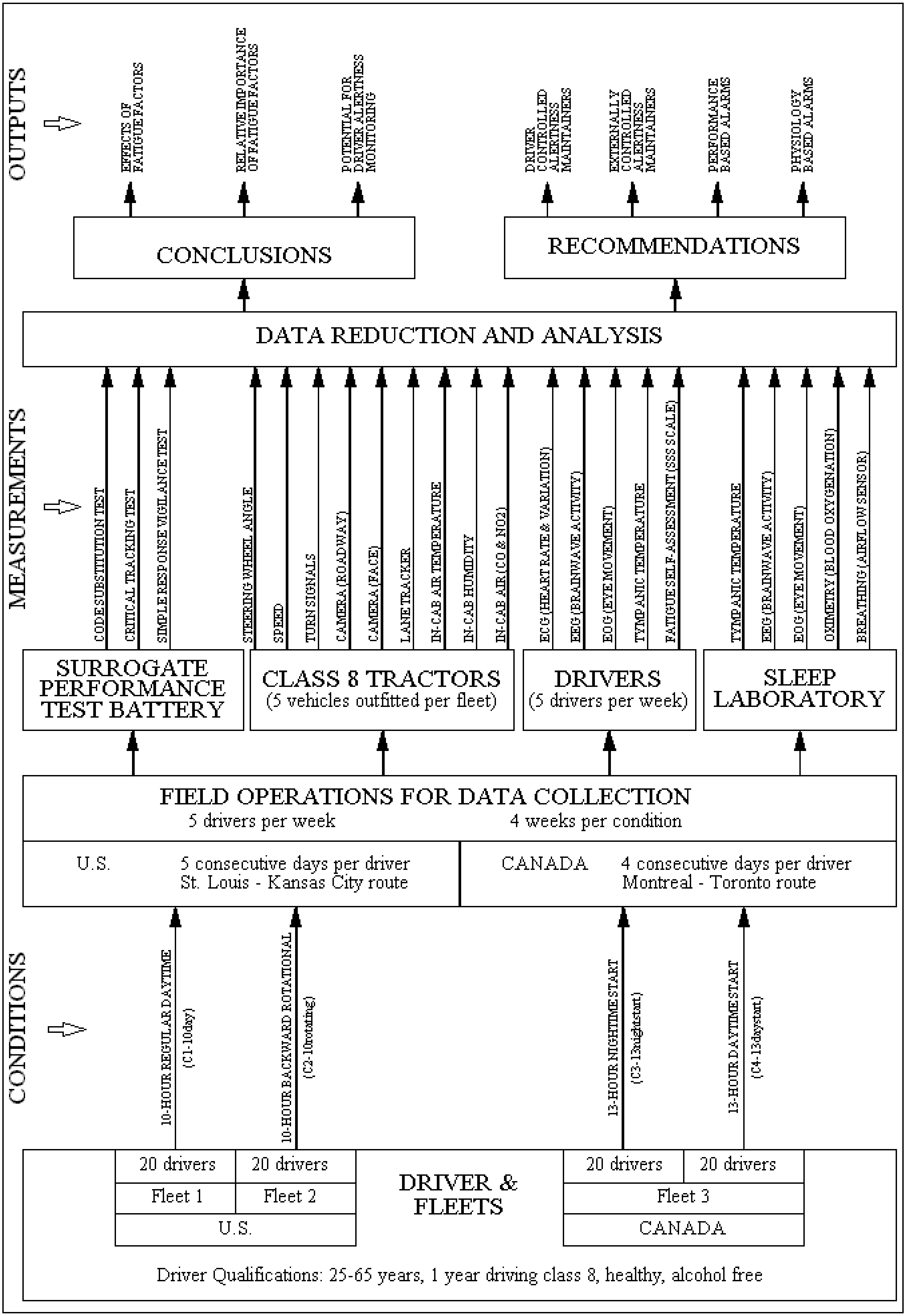

Figure 1. Overall data collection program and study design

Table 1. Comparison of U.S. and Canadian hours-of-service rules relevant to this study

HOURS-OF-SERVICE RULES	U.S.	Canada
Driving time limitation	10 hours	13 hours
On-duty time limitation	15 hours	15 hours
Off-duty time minimum	8 hours	8 hours
7-day on-duty time limitation	60 hours	60 hours

DRIVING SCHEDULES

The driving schedules employed in the data collection are summarized in Table 2 and, in this study, are identified specifically as observational Conditions C1-10days, C2-10rotating, C3-13nightstart, and C4-13daystart. Each of the four driving Conditions involved a different group of 20 drivers. Table 2 also presents the calculated percentages of night (0000 to 0600) driving time for the outbound and inbound directions of each observational Condition to provide a time of day (circadian) perspective of each for comparison purposes.

The "baseline" and "rotating" Conditions (C1-10day and C2-10rotating) involved 10 hours of driving under U.S. hours-of-service rules. Each of the drivers in these two Conditions repeated this trip five consecutive times. In the "baseline" Condition (C1-10day), the drivers started their trips at about the same time each morning. The round trip involved just short of 10 hours of driving, and about 12 hours on duty. This represented the most benign schedule of the four that were employed. In the rotating Condition (C2-10rotating), we achieved about a 3-hour setback, with the first trip's departure time (on average) at 1000, the second trip's departure time at 0730, the third departure time at 0500, the fourth departure time at 0130, and the fifth departure time at 2100 of the same day.

Condition C3-13 nightstart consisted of four consecutive trips requiring 13 hours of driving and 15 hours on duty, leaving Toronto about 2300 (on average), dropping a trailer and picking up another in Montreal, and returning to Toronto. Condition C4-13daystart covered the same route as Condition C3-13nightstart, but involved a different set of 20 drivers who departed from Montreal at about 1300 (on average), dropping and hooking in Toronto, and returning to Montreal for four consecutive trips.

For each condition, five drivers were scheduled for trips on each day and they were nominally scheduled to depart at 45-minute intervals to allow field personnel time to process each one properly.

ROUTES DRIVEN

The U.S. drivers' route took them from St. Louis, Missouri, to the metropolitan area of Kansas City Missouri - Kansas. The distance between St. Louis (metro population 2.4 million) and Kansas City (metro population 1.5 million) is approximately 400 kilometers (250 miles). The route between the origin and destination terminals consists primarily of interstate highway driving, but with typical urban traffic congestion within the urbanized areas of the two cities. The data collection in the U.S. involved two motor carriers and occurred over the period June 14 to August 22, 1993. Over an 8-week period, the 40 drivers, divided into groups of five drivers per week, each drove five consecutive round trips.

Table 2. Driving schedules used as study "Conditions" including associated percentages of night (0000 to 0600) driving time

1. **Condition C1-10day, "10-hour daytime":** 20 drivers operated on a 10-driving-hours turnaround route, starting at about the same time each morning (about 1000 on average) for five consecutive trips. Average percentages of night (0000 to 0600) driving time during outbound and inbound trips were 1.6% and 2.4%, respectively.

2. **Condition C2-10rotating, "10-hour rotating":** 20 drivers operated on a 10-driving-hours turnaround route, starting about 3 hours earlier each day (initial trip about 1000 on average) for five consecutive trips. Average percentages of night (0000 to 0600) driving time during outbound and inbound trips were 17.7% and 4.4%, respectively.

3. **Condition C3-13nightstart, "13-hour nighttime starts":** 20 drivers operated on a 13-driving-hours turnaround route, starting at approximately the same time, late each evening (about 2300 on average), for four consecutive trips. Average percentages of night (0000 to 0600) driving time during outbound and inbound trips were 68.1% and 5.4%, respectively.

4. **Condition C4-13daystart, "13-hour daytime starts":** 20 drivers operated on a 13-driving-hours turnaround route, starting at about the same time in the late morning and early afternoon (about 1300 on average), for four consecutive trips. Average percentages of night (0000 to 0600) driving time during outbound and inbound trips were 2.8% and 38.2%, respectively.

The Canadian over-the-road portion of the data collection included 40 drivers from one motor carrier. Over the 8-week field data collection period, five drivers per week each drove four round trips over the Montreal-Toronto route. The Condition C3-13nightstart drivers departed their Toronto terminal in the late evening, while the Condition C4-13daystart drivers departed from their Montreal terminal during the late morning/early afternoon. The distance between Montreal (metro population 3.1 million) and Toronto (metro population 3.9 million) is 530 kilometers (331 miles). Similar to the U.S. route, the route between the Canadian origin and destination terminals consisted of highway driving, but with congestion within the limits of the two cities. The data collection in Canada was completed over the period September 27 to December 3, 1993.

VEHICLES

The truck tractors used in the data collection were from each volunteer motor carrier's line-haul fleet, and were late-model units that were dedicated to this project for the duration of the data collection at the host's operation. The trucks were maintained by the motor carriers' maintenance personnel during the data collection.

Fleet participation was sequential. Five conventional Class 8 tractors from each participating motor carrier were outfitted with on-board monitoring equipment. In Conditions C1-10day and C2-10rotating, the first two motor carriers' standard configuration of a single-drive-axle tractor and twin 28-foot trailers was used, with the tractors being short-nose conventionals for fleet 1 and long-nose conventionals for fleet 2. In Conditions C3-13nightstart and C4-13daystart, the third motor carrier's standard configuration of a long-nose conventional, tandem-drive-axle tractor and single 45, 48 or 53-foot semitrailer was used. Since the same carrier sponsored both Conditions C3-13nightstart and C4-13daystart, the same five tractors were used throughout. All participating drivers were completely familiar with their respective assigned vehicles.

It was not possible to use the same configuration of vehicles throughout the study to control unwanted variation. Any variance introduced by using different configurations would primarily affect between-Condition comparisons rather than within-Condition comparisons. Based on anecdotal reports, track test results reported in the literature (MTO, 1996a; MTO, 1996b; FHWA, 1996a) and task analyses, we believe the variation introduced by the vehicle differences was of secondary importance in this study, which primarily involved open highway driving on relatively straight and flat terrain.

DRIVER QUALIFICATIONS, RECRUITMENT, AND TRAINING

Eighty properly qualified commercial drivers served in the various driving Conditions of the data collection program. All drivers were subject to common selection and qualification procedures.

The drivers were all male and were between the ages of 25 and 65 years. Drivers had to have at least one year of experience driving Class 8 tractors, to be healthy and free from controlled substances and alcohol, and to have no documented medical history of sleep disorders. Each driver was given an examination by a physician, including urinalysis and drug screening. We had intended to randomly select 20 drivers from a much larger pool, but when we imposed our uniform age distribution requirement, we found that 20 drivers comprised virtually 100% of the available pool at each site.

Driver Selection

The selection process began with a group briefing of drivers by the principal investigator and representatives from FHWA and TRI, and in Canada by Transport Canada. All drivers were volunteers. The drivers had to be willing to drive the specified routes and schedules and to sleep in the sleep-laboratory quarters provided. They also had to permit the ECG, EEG, and EOG hookups and to submit to scheduled and random alcohol and drug testing.

Driver Qualifications

DOT Qualification: The motor carrier employing each driver who participated in the study certified that the driver was qualified per 49 CFR 383 and 391. These regulations govern licensing, minimum age, language skills, vehicle operating skills, physical qualifications, and freedom from disqualifying offenses.

Health: Each carrier certified that its participating drivers had met the physical qualifications of 49 CFR 391.41, as evidenced by a medical examiner's certificate per 49 CFR 391.43. This exam is required every two years, but if a driver's exam had not been administered within the prior four months, we required another examination before he participated.

Age: We attempted to achieve a uniform distribution of ages in the ranges 25-35, 36-45, 46-55, 56-and-up in each of the four groups of 20 drivers who served on the four different schedule Conditions. The groups passed the Bartlett test for homogeneity of group variances (CHI-SQUARE = 5.494, df = 3, p = 0.139), but ANOVA indicated different group means, F(3,76) = 5.221, p = 0.002. The group mean ages were: Group 1 = 49.0, Group 2 = 43.8, Group 3 = 40.3, and Group 4 = 38.0. Tukey post-hoc comparisons showed that Group 1 was older than Group 4 (p = 0.002) and older than Group 3 (p = 0.023).

Gender: Beilock's survey (1990) sample indicated that females made up about 3% of the driver population, so women would have been slightly overrepresented if we included even one woman in each observational Condition of 20 drivers. Furthermore, no statistically significant conclusions could have been drawn concerning female drivers unless many per Condition were included. Consequently, only male drivers were selected to serve in the data collection.

Experience: Each driver was required to have had at least one year of current experience operating heavy commercial vehicles of the type he was expected to drive during the data collection.

Controlled Substances and Alcohol: The drivers participating in the study were all subject to prohibitions against driving under the influence of drugs or alcohol contained within their respective U.S. (e.g., 49 CFR 392.4 & 392.5) and Canadian government regulations, as well as any additional requirements imposed by their employers. A urinalysis test was required for each driver twice during the week spent as a subject for the investigation. The first sample was acquired prior to the first major sleep period of the week. The second sample was acquired prior to the second through fifth major sleep periods, with random assignment and with no prior announcement. We also administered breath analyzer tests to each driver at the beginning of each trip. Urinalysis and breathalyzer tests were performed only as a screen to safeguard the research protocol, by permitting assessment of potential data outliers from these perspectives. No driver’s data had to be removed as a result of these tests.

Informed Consent: The study design was developed to comply with existing hours-of-service regulations. The drivers who participated in the data collection were informed that their driving schedules might be very demanding, involving many hours of driving and on-duty, up to the limits of the regulations. Each driver was advised that he was free to pull over and take a rest break any time he felt it was in the interest of safety, and that he was free to withdraw from the data collection entirely if he wished to do so for whatever reason. After reading the information on the Informed Consent Form, each driver documented his informed consent by signing the form, before being allowed to participate in the data collection. The Informed Consent Form was reviewed and approved by FHWA and Transport Canada.

Driver Duties

Each driver's principal duty was to drive his instrumented truck tractor and trailer(s), loaded with freight, from the starting point to the turnaround point, then to drive back to the starting point with another trailer(s) and load of freight. None of the drivers was required to load or unload cargo as part of their normal duties, and they did not do so during the study. Fleet practices at some locations required drivers to drop and/or hook their own trailers (e.g., couple and uncouple). This task was not physically demanding or time consuming in and of itself, although locating the designated trailer was sometimes time consuming.

Driver Training

Prior to his participation in the field testing, each driver was required to participate in three 1½-hour training sessions, on three separate days, during the week preceding his assigned data collection period. Drivers practiced the three performance tests during these sessions in order to minimize practice effects during data collection.

Driver Compensation

All data were collected during the course of revenue-generating trips. The drivers were paid their regular salaries by their respective motor carrier. Because the data collection protocol required the drivers to spend each night in designated away-from-home sleeping quarters, drivers were paid their customary compensation for layovers out of project funds. Likewise, the protocol required the drivers to spend a total of about one and one-quarter hours per trip to take their performance tests. The drivers were reimbursed for this activity at their usual hourly rate, out of project funds.

MEASUREMENT OF DRIVING TASK PERFORMANCE

For the purposes of this research, we were concerned with the drivers' proficiency at tasks related to safely controlling moving vehicles. Because the data were collected in an actual operating environment, other tasks, such as hazard recognition, evasive maneuvering, and skid control/recovery, were not included in the study design. In the following subsections, we will discuss the direct measures of driving task performance that we used in this study.

LANE TRACKING

Because proficiency at keeping the vehicle in its lane has obvious safety relevance, and because prior research indicated that performance of this task is sensitive to fatigue, driver lane tracking performance was monitored throughout the study. Human Factors Research, Inc. designed and constructed an automatic lane tracker which permitted much greater precision in measuring lateral lane position than is possible with on-board observers (Mackie and Miller, 1978). Six copies of this device were fabricated for use in the present study.

The lane tracker electronically locks on to the lane delineation (e.g., painted line). The vehicle's position in the traffic lane is measured by electro-optically scanning a section of road surface about 3.05 meters (10 feet) in front of the vehicle, perpendicular to the direction of travel. The lane tracker is housed in an environmental enclosure external to the tractor cab, just above and to the right of the windshield (see Figure 2).

Figure 2. Lane tracker mounted on roof of truck cab

The 256-element charge-coupled device (CCD) array then allows resolution of the vehicle's lateral position to less than 12.7 mm (0.5 in.). The vehicle's lane position was sampled by the microcomputer five times per second and recorded for later analysis. At night, the lane line was illuminated via an exterior lamp which was louvered to prevent glare.

When the lane tracker data editing was completed (e.g., to remove periods of time during which the lane tracker had lost track of the lane), the remaining data were processed into one-minute epochs and then aggregated further into 10-minute epochs based on a pre-defined methodology that ensured a valid basis for analysis. Average lane tracker coverage across the four observational Conditions was 33%. Coverage was affected by several factors, including variation in the contrast of the painted line against the background of the pavement (faded and painted-over lines), the weather (rain produced a reflectivity that reduced the visibility of the edge line), widely varying conditions of natural illumination during the day (sunlight at an angle that produced glare in the region scanned by the lane tracker) and artificial illumination at night. However, based on the knowledge of the optoelectronic functioning of the lane tracker, there was still considered to be a sufficient sample to draw sound conclusions. Furthermore, there was no reason to expect that a smaller or larger sample of lane tracking would have affected the lane tracking standard deviation statistic. This is supported by the fact that the correlation between lane tracking standard deviation and the number of analyzable samples in each 10-minute epoch was not statistically significant (Pearson $r = -0.01$, $p = 0.14$).

STEERING WHEEL MOVEMENT

Steering wheel movement had long been a favored measure of driver behavior among researchers (Harris and Mackie, 1972) due to its obvious relevance to vehicle control and the ease with which it can be recorded.

Steering wheel movement was acquired using a precision, continuous turn, servo potentiometer coupled to the steering column by a gear formed by fastening a toothed belt around the steering column. The instrumentation computer recorded steering wheel angle 20 times/second. The steering wheel movement magnitude data have approximately one degree resolution.

SPEED AND DISTANCE

Speed variability is known not to be as good an indicator of fatigue as lane tracking variability. Moreover, most of the vehicles in this study were speed governed. Thus, speed variability was not used as a measure of fatigue. However, vehicle speed and distance data were available to this study and they permitted automated removal of low-speed data segments from speed sensitive vehicle control data such as steering and lane tracking, facilitated analysis and interpretation of driving EEG data by precisely identifying where and when the vehicle was being operated, and allowed compensation for road geometry during analysis of lane tracking and steering data by permitting identification of precise locations along the route.

Odometer data were acquired from the electronic speedometer systems already existing in the vehicles. Pulses were either picked directly off the transmission's tachogenerator output or off the signal used to engage the solenoid driving the vehicle's mechanical odometer.

MEASUREMENT OF PERFORMANCE ON SURROGATE TESTS

We know that surrogate tests can be used to measure human capabilities representative of the level of proficiency in the performance of driving tasks. Appropriately selected surrogate performance tests require the driver to do something which can be expected to utilize the same psychoneurological processes as those necessary for the performance of the real operational tasks of driving.

The off-line performance tasks included in the test battery for this study were the Code Substitution Test (CS), the Critical Tracking Test (CTT), and a variant of Wilkinson and Houghton's (1982) Unprepared Simple Reaction Time test which we will refer to as a Simple Response Vigilance Test (SRVT). These tests were selected because: they have previously been used in other operator fatigue and impairment studies; they would provide a sufficiently broad range of psychomotor and cognitive performance measures; they could be packaged and performed using the "main" data collection microcomputer on board the vehicle; and they would take twenty minutes or less time, in total, per administration.

CODE SUBSTITUTION (CS) TEST

This test requires the driver to press a numbered key that corresponds to the number assigned to a displayed letter on the screen, using the code of relationships between letters and numbers also displayed on the screen. This test measures information processing, perceptual speed, rapid visual search, and response selection speed.

CRITICAL TRACKING TEST (CTT)

The CTT is a highly reliable and accurate test of psychomotor control. It uses carefully developed, complex algorithms to move a pointer in an unpredictable fashion along a horizontal axis on a computer display. The driver's task is to move the steering wheel, which is linked to the test computer by a gear and servo potentiometer, such as to keep the pointer at the indicated center of the display. The task becomes increasingly demanding during the course of each 30-second presentation. The difficulty of the task at the point which the subject fails it reflects perceptual-motor response time between eye and hand. The CTT has been shown in other studies to be sensitive to driving time and time of day (e.g., circadian effects) (Mackie and Miller, 1978).

SIMPLE RESPONSE VIGILANCE TEST (SRVT)

The Simple Response Vigilance Test (SRVT) used in this study is remarkably simple in its implementation. In each trial a visual stimulus in the form of three digits ("000") is displayed. This is essentially a timer which begins counting immediately after it is made visible to the subject. The subject simply presses a response button as quickly as possible after the stimulus appears. After response, the reaction time figures on the counter are displayed for 1.5 seconds before they disappear. After a variable interval, from 1 to 10 seconds, the stimulus again appears and a new trial begins. The preferred duration of this test is 10 minutes according to Wilkinson and Houghton (1982), and that was the test length used. Wilkinson and Houghton claim that few tests that are sensitive to the effects of stress will have such a short practice effect. They cite evidence showing the test to be sensitive to the effects of many of the stressors that normally influence arousal. This test measures the ability of a person to remain alert for the relatively rare, semi-random occurrence of a repetitive stimulus in a boring environment.

TEST BATTERY ADMINISTRATION

The performance test battery was self-administered, by each driver, four times daily for the 10-hour-driving group, and three times daily for the 13-hour-driving group (a fourth administration had to be eliminated because it would have caused these drivers to exceed allowable on-duty time limits). These tests were undertaken prior to the commencement of each trip, at the mid-trip turnaround point, and again on arrival at the trip end-point. Drivers participating in the 10-hour-driving schedules (Conditions C1-10day and C2-10rotating) also performed the tests just prior to departure from the mid-trip turnaround point. This was done to permit measurement of any recovery that might have taken place as a function of the mid-trip break.

There was no rest-time between the three tasks in the battery. The Code Substitution (CS) test took three minutes, the Simple Response Vigilance Test (SRVT) took ten minutes, and the total time for the five Critical Tracking Test (CTT) trials varied between three and five minutes. With administrative time (positioning keyboard and push-button, checking steering wheel position, entering his personal identifier, paging through instruction pages) the complete test battery took approximately 18 minutes. Each successive test session presented the tests in the battery in a different and predetermined order, in order to minimize order effects.

The tests ran on the on-board instrumentation computer. Tests presented visual stimuli to drivers via a 9" VGA monitor, recorded driver responses, calculated measures summarizing driver performance, and stored the data for later downloading. Data were recorded as printable ASCII text files by the truck instrumentation and transferred to 3.5" diskettes each day.

VIDEO RECORDING OF DRIVER'S FACE AND ROAD AHEAD

Throughout each drive, simultaneous video records were made of each driver's face and of the view of the roadway ahead. A forward-looking camera was used to aid in evaluating critical roadway-oriented events. This gave the researchers the ability to view the lateral position of the vehicle in the roadway as well as a dynamic view of road and traffic conditions. The camera was mounted inside the cab, looking forward through the windshield, with approximately the same view as the vehicle driver. The camera had sufficient light sensitivity for nighttime use. A second camera was mounted on the dash to provide a view of the driver's face. One infrared illuminator provided adequate near-IR illumination which was visible to the camera but not to the driver (see Figure 3). The view of the road ahead was inserted electronically into part of the field of view of the camera recording the driver's face (e.g., the lower 1/3 of the picture).

Figure 3. Face camera and infrared illuminator

Because of the large volume of data, each video tape was sampled for several minutes at approximately half-hour intervals of trip time. For each video sample, reviewers trained to make the required judgments completed a 50-item form (e.g., the Video Detail Record Form) regarding various aspects of the visual scene down the road and of the driver's face. One of the judgments the video reviewers made and recorded on the form was whether the driver "appears drowsy." The researchers had used this technique before with good results, to judge operator monitoring in simulated sonar operations (Wylie et al., 1988). Wierwille et al. (1994, 1995) conducted tests of driver drowsiness judgment from driver video, and reported that the results were valid and reliable.

The principal investigator spent about one hour with the video reviewers in a group, explaining the desired procedures for sampling the video, viewing it, and recording the results.

Particular emphasis was placed on the judgment "appears drowsy" and several different video examples were viewed together until the research assistants were responding alike. Emphasis was also placed on the importance of partially closed and slowly closing eyelids as a cue for recording the judgment "appears drowsy" on the form. Among others, cues used in establishing a drowsy judgment included behavioral patterns reflected by "droopy eyelids," "slow blinks," "head nods," "eyes shut for ___ seconds," and "yawns." The research assistants were periodically monitored by the principal investigator during the course of sampling the video. From reviewing these judgments, it can be said with confidence that the drivers who were judged "drowsy" ranged from slightly drowsy to extremely drowsy. Most frequently, the "drowsy" driver was somewhere in between these extremes.

Sampling the videos every 30 minutes (one out of every five 6-minute epochs) provided an unbiased estimate of prevalence of drowsiness, but revealed only limited information on duration of drowsiness. Duration of drowsiness (e.g., number of consecutive 6-minute epochs judged "drowsy") was of interest for basic physiological and ergonomic reasons, and also because longer periods of driver drowsiness (and/or greater numbers of short periods of drowsiness) were hypothesized to be associated with greater exposure to risk, since each additional 6-minute epoch judged "drowsy" is associated with a separate and unique place-time-driver combination of factors that might lead to a crash. Although no crashes occurred, that outcome does not provide a basis to accept or reject the hypothesis.

Therefore, when the video database sampling at 30-minute intervals was complete, a second phase of review was undertaken, in which each instance previously noted "drowsy" was viewed continuously from 30 minutes before to 30 minutes after its occurrence. A 50-item Video Detail Record Form was completed for each of the ten 6-minute epochs in the review of each drowsy event detected by sampling.

PHYSIOLOGICAL MEASURES

BODY TEMPERATURE

Body temperature was estimated periodically (usually bi-hourly) during waking using tympanic membrane temperature as determined by an infrared ear probe (Thermoscan Inc, San Diego CA). To acquire all of the measurements needed, we developed a procedure and taught the drivers to take and record their own temperatures.

There are many circadian rhythms in the body. Measuring only body temperature provides a limited view of the degree of synchronization of the body with external time cues. The single measurement gives no information about internal dissociations among various rhythms in the body. However, by acquiring the single measure, we were in a position to detect instances when a group of drivers in the project was not synchronized with the day-night cycle.

POLYSOMNOGRAPHY (PSG) DURING SLEEP

The study used the most widely accepted technique in the sleep research community to measure sleep, and hence calculate the amount of sleep obtained. This was by means of continuous EEG measurement done according to the methods defined by Rechtschaffen and Kales (Rechtschaffen & Kales, 1968) and performed within the context of polysomnography (Guilleminault, 1982). The measurement was required because, in any study of worker fatigue, it is necessary to document not only the time of day that work takes place but also the amount of sleep obtained preceding each work period. In addition to the well-known circadian effects, and particularly of the increasing sleep tendency during the typical sleep period from about 11 p.m. to 7 a.m. (Broughton, 1975; Richardson et al, 1982), the amount of sleep a person has obtained in the preceding 24 - 48 hours is the most important determinant of sleep tendency during a period of intended wakefulness (Carskadon & Dement, 1981; Rosenthal et al, 1993).

During this study, all subjects slept in comfortable rooms located as near as possible to the travel route under study. Subjects underwent their initial hook-up, before their first major sleep period in the laboratory, in a dedicated room adjacent to the sleep rooms about 60-90 minutes before retiring. Oxford tin electrodes were used to record central and occipital EEG, lateral EOG (outer canthi), chin EMG (submental muscle), respiratory airflow (nasal thermistor), respiratory effort, and oxygen saturation of arterial blood (SaO_2; using a finger probe). Hook-ups for the principal sleep period subsequent to the first trip consisted of replacing loose or poorly functioning leads and repositioning sensors for chest motion, air flow and SaO_2. The recording montages are shown in Table 3. The electrode placements (channels 1-5) allowed signal acquisition that supported the standardized polysomnographic scoring method (Rechtschaffen and Kales, 1968). During sleep, channels 5-7, in conjunction with fingertip oximetry, allowed assessments of sleep apnea.

Each EEG/EOG/EMG record was reviewed by a trained polysomnographic technologist. Scoring was performed with 30-second epochs according to the standard method of Rechtschaffen and Kales (1968). Scoring consistency among the various technologists was assured by re-scoring randomly selected records, group discussions of difficult records and review of data by a diplomate of the American Board of Sleep Medicine. The allowable EEG scores were as shown in Table 4.

Oximeter output was screened by the technologist for significant signs of sleep apnea. Drivers with pronounced respiratory abnormalities during sleep were reported to one of the investigators who, in turn, notified the driver of the abnormal finding and suggested privately to the driver that the driver seek medical evaluation. If a driver was found to have clinically significant sleep apnea, specialists in apnea scored all of that driver's sleep records.

Table 3. Electrode and sensor placements for sleep and driving
The central (C), auricular (A) and occipital (O) electrode sites were as described in the International 10-20 System (Jasper, 1958). The left-outer-canthus-right-outer-canthus (LOC-ROC) derivation was described, for example, by Osselton (1974).

Channel	Sleep	Driving
1	C_4-A_1	C_4-A_1
2	O_1-A_2	O_1-A_2
3	LOC-ROC	LOC-A_2
4	C_3-A_2	ROC-A_1
5	EMG-EMG (chin)	C_3-A_2
6	Chest motion	---
7	Oral/nasal airflow	---
8	ECG	---
Indifferent	C_z	C_z

Table 4. EEG scores assigned manually by polysomnographers, as specified by Rechtschaffen and Kales (1968)

Score	Meaning
M	Motion artifact; unusable signal
W	Awake; usable signal
1	Stage 1
2	Stage 2
3	Stage 3
4	Stage 4
R	Stage REM

POLYSOMNOGRAPHY (PSG) DURING DRIVING

We identified periods during which the truck was moving and the polysomnographer had scored the EEG, according to Rechtschaffen and Kales methods, as showing signs of sleep (Stages 1, 2, 3, 4, or REM). This analysis was triggered by the observation by O'Hanlon and Kelley (1977) of obvious stage 2 sleep and delta activity in the EEGs of two drivers on the highway. We wanted to know whether or not there were similar incidents in our database.

After lights on, at the end of a principal sleep period, sensors for air flow and SaO_2 were removed and the remaining leads were checked for stability and electrical impedance. The montage connections were changed to produce the channels shown in Table 3 for the subsequent driving period.

Epoch lengths of 20 seconds were used for the waking EEG analysis. This epoch length was suggested by the EEG analysis results of Mackie and Miller (1978). This is also an alternative epoch length used in clinical sleep scoring. Each 20-second record's human scorer information concerning the Rechtschaffen and Kales (1968) sleep stage (Table 4) was time-scaled against the segments during which the truck speed equaled or exceeded 45 mph.

Our technologists were aware that the driving EEG data came from a driver's work period and that some naps had been taken with the truck parked. By intent, however, they were unable to determine during the scoring process whether or not the truck was moving. When naps were flagged by the field personnel, the scoring technologists were still unsure of the exact time boundaries of the naps. Thus, the technologists were blinded to truck speed for nearly all of their scoring. This blinding was a planned element of initial scoring that prevented the technologists from being improperly influenced by knowledge of vehicle motion.

QUANTITATIVE EEG (QEEG) DURING DRIVING

In addition to the PSG analyses of the EEG data collected during driving, quantitative EEG (QEEG) analyses were also performed. Because of the developmental nature of this work, neither conclusions nor recommendations in this report were based upon the QEEG analysis. These analyses and the respective results and conclusions are, however, described in an appendix to the study report for reference in future research.

VAGAL TONE

During driving, the lead-II-appearing electrocardiagram (ECG) was acquired from the CR-5 lead (Blackburn, 1969; Simonson, 1971) by the Vagal Tone Monitor (Instrumentation for Medicine, Greenwich, CT). Self-adhesive spot electrodes were used that are also used for treadmill exercise testing.

At the time of writing this report, the vagal tone data remained in raw form and had not been fully examined for the following reasons. First, the vagal tone data were viewed as having lower potential value than other physiological measures for the overall investigation. Many data were lost due to electrode failures in the field. Second, we found through other unrelated investigations that occurred after the data collection in this project had begun, that the vagal tone monitor had greater value in discriminating brief behavioral arousal from a background of low arousal than in discriminating among various levels of low arousal (Miller et al., 1992; Miller, 1994). In this respect, it was somewhat similar to the electrodermal response. We felt that the PSG and QEEG physiological data would allow better discrimination of shifts, for example, from relaxed wakefulness into drowsiness. Finally, the extent of the work required for both QEEG and vagal tone monitor data reduction and analysis exceeded the estimated effort allocated to it in the project by a large amount.

DRIVER-SUPPLIED INFORMATION

DRIVER QUESTIONNAIRES AND LOGS

Each driver, on first reporting to the sleep laboratory, filled out a detailed "sleep history questionnaire" containing seventy questions concerning sleep history, habits and medical condition. He filled out a questionnaire prior to and at the end of each rest period during the data collection. A daily log was also kept concerning each day's activities on the road, including stops, road and weather conditions, noteworthy driving events, and meals taken. Each driver recorded his tympanic temperature measurements on a log sheet provided to him for this purpose.

SELF-ASSESSMENT OF FATIGUE

Numerous scales have been developed and successfully used for self-assessment in a variety of contexts including studies of heavy vehicle driver fatigue (e.g., Mackie and Miller, 1978). We used the Stanford Sleepiness Scale (Hoddes et al., 1973) presented in Table 5.

Table 5. Stanford Sleepiness Scale ratings

Rating	Definition
1	Feeling active and vital; alert; wide awake
2	Functioning at high level, but not at peak
3	Relaxed; not at full alertness; responsive
4	A little foggy; not at peak; let down
5	Fogginess; losing interest in staying awake; slowed down
6	Sleepiness; prefer to be lying down
7	Almost in reverie; sleep onset soon; hard to stay awake

Each driver assessed his level of fatigue according to ratings on the Stanford Sleepiness Scale five times over 24 hours, throughout each driver's participation in the study: before and after his principal sleep period; beginning of trip; arrival at, and departing from, mid-trip turnaround terminal; and, end of trip.

TRUCK CAB ENVIRONMENT

HEAT, NOISE, AND VIBRATION

Heat is known to affect driver performance (Mackie, O'Hanlon, and McCauley, 1974). Dry-bulb air temperature, globe temperature, and relative humidity data were collected with Omega

Model HX93 sensors interfaced directly to the instrumentation computer. The on-board computer calculated wet bulb globe temperature, the NIOSH de facto standard for heat stress.

Mackie et al. (1974) also showed that noise and vibration are much less consequential to driving performance, though not necessarily to long-term driver health. They concluded that their results "generally confirm the fact that drivers in suitably designed seats, air-cushioned or otherwise, will infrequently experience vibration of a magnitude above the presently recommended 8-hour boundary for `reduced performance efficiency." Noise and vibration data were not recorded.

AIR QUALITY

Ziskind et al. (1977) concluded from their study of toxic gasses in heavy duty diesel truck cabs that "NO_2 levels are sufficiently elevated in many vehicles to be of concern" (page 5). While the study team believes that this concern has more to do with longer-term occupational health than with immediate impact on driver performance, a decision was made to monitor exposure.

Personal daily-use Leak-Tec paper dosimeter badges were used to monitor NO_2 and CO concentrations in each truck cab. The badges use replaceable indicating papers which change color according to the concentration of gas present.

ON-BOARD INSTRUMENTATION

DATA ACQUISITION COMPUTER

An on-board microcomputer was used to record data generated by the lane tracker, the steering wheel angle sensor, and other instrumentation. The microcomputer allowed the off-line performance tests (CS, CTT, SRVT) to be done in the cab, eliminating a possible "alerting" effect from a driver walking from the cab to the field laboratory. It was also used to provide a means to signal a driver if he were to get underway without having his ECG cable plugged into the instrumentation. A schematic of the on-board data collection systems is presented in Figure 4.

INSTRUMENTATION ENCLOSURE

Most of the instrumentation was housed in a transportation enclosure designed for electronic equipment. This was mounted in place of the passenger's seat in each of the five tractors provided by each of the participating motor carriers (see Figure 5). A sixth complete unit was built and used as a spare. The enclosed equipment included a rack-mount blower unit for cooling, sensor interface box on slides, transformer mounting panel, instrumentation computer, vagal tone computer, and video shelf to hold VCR, video splitter, and modem for time code recording. The total unit weighed about

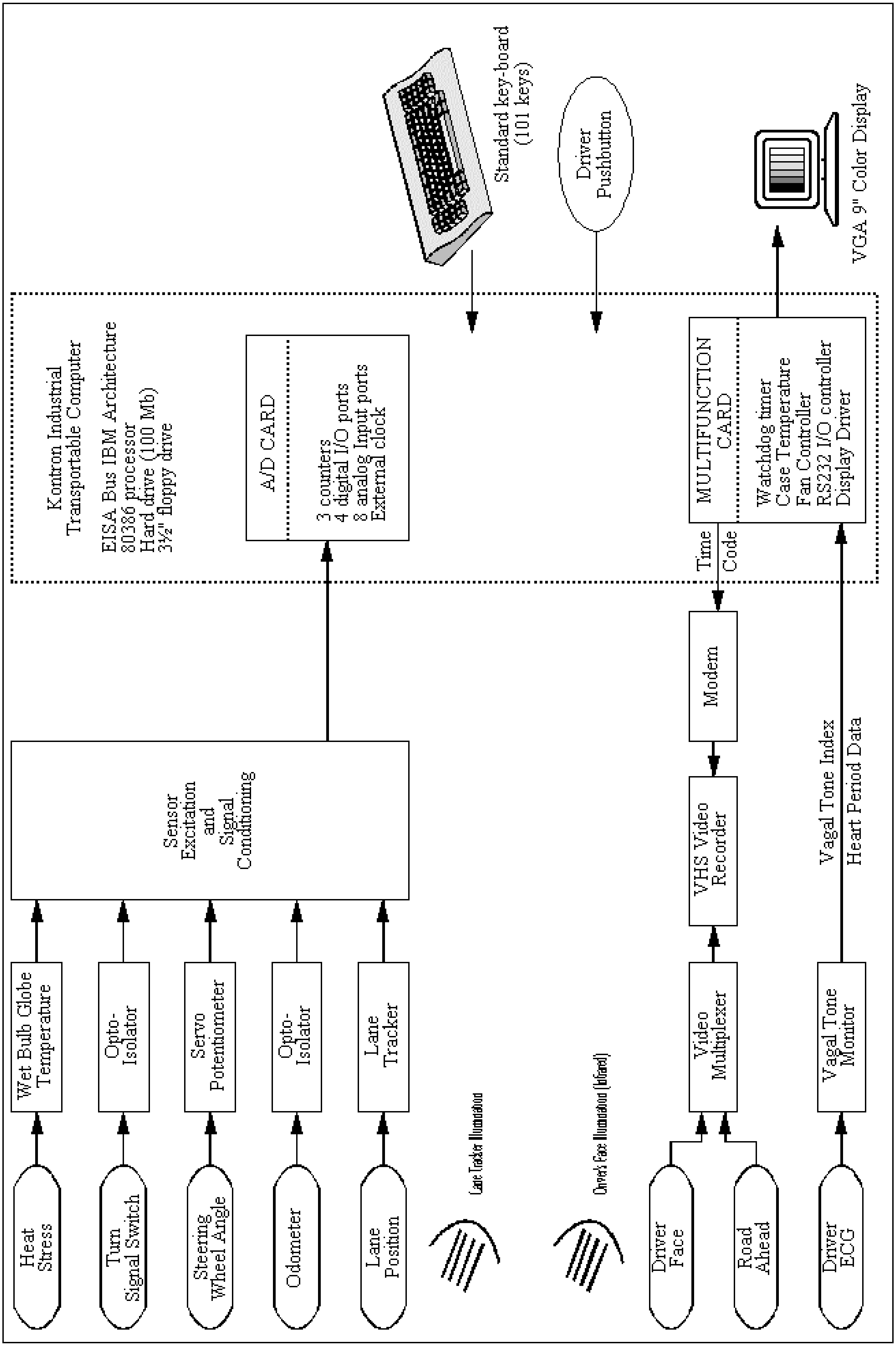

Figure 4. On-board instrumentation schematic

82 kg. (180 lbs.) and was held to its mounting platform on the truck floor by six elastomeric mounts to attenuate shock and vibration.

ELECTRICAL POWER SUPPLY

Electrical power for the instrumentation was provided by a Honda EM1600X portable gasoline-engine generator mounted on the outside of each truck tractor's right-hand frame rail behind the cab.

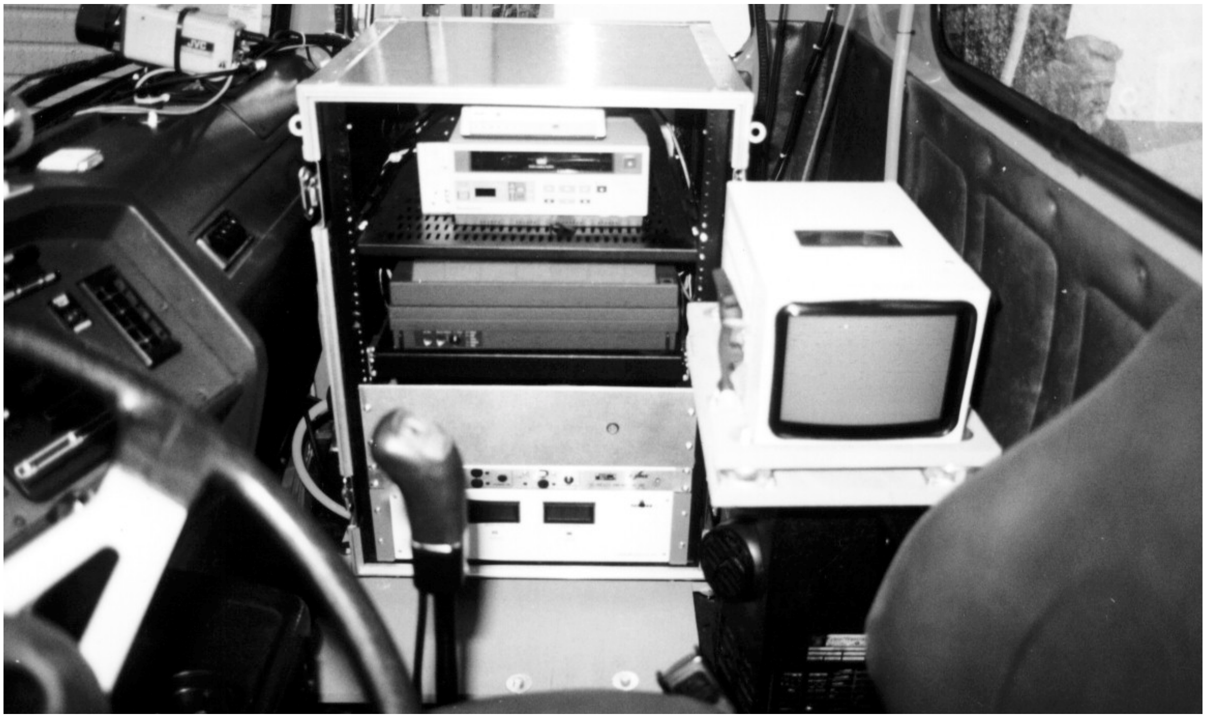

Figure 5. Instrumentation in cab

FIELD CONTROL PROCEDURES TO ENSURE DATA QUALITY

Because of the complexity of the equipment and procedures proposed for driver and vehicle data collection and analysis, a pretest was conducted over a 5-day period in July 1992 to demonstrate the integrated functioning of the equipment, and to ensure that data collection and analysis procedures were able to provide the appropriate level of quality.

As we proceeded to collect data in each one of the four observational Conditions, we moved our base of operations in turn to each of the four terminals designated by the participating motor carriers. We installed the instrumentation in each of five truck tractors that the host motor carrier had dedicated to the data collection, and started training drivers in the performance tests. We hired local field technical support (six research assistants and two electronic technicians) on a temporary basis.

Departure, turnaround, arrival, and data review checklists were prepared and followed for each driver and trip. These also included performing trouble shooting and hardware/software maintenance as necessary throughout the period of time available until the trucks had to be back on the road for the succeeding trip. The checklists comprised an extremely valuable aspect of the data documentation, notations on them often aiding to interpret questions about various segments of data.

THE DATABASE

The database collected during the field operation covers over 200,000 miles of driving. It includes some 4,000 hours of video data, 9,000 hours of physiological recordings (heart rate, eye movement, breathing, and brain wave), and 700 megabytes of real-time truck computer digital data records. In the study reported here and involving four data collection Conditions, there were 360 round trips and 400 principal sleep periods. Data quality was good, especially considering the challenging conditions of a truck traveling 800 kilometers (500 miles) per day in the U.S. and more than 1,160 kilometers (662 miles in Canada). The research team estimates that the size of the summary data base will be between 1,600 and 2,400 megabytes.

Table 6 describes the collected raw database in summary form. The reduced data set that was produced from the raw data as described in this chapter and used in the analysis is described in Table 7.

Table 6. The raw database collected from the field operations

DATA SOURCE	MEDIA TYPE	NUMBER OF MEDIA	AMOUNT OF DATA
Video Recordings	VHS Cassettes: T-120 (USA); T-160 (Canada)	720	4,000 hours
Truck Computer Recordings	3.5" HD diskettes	500	700 Mbytes
Physiological Recordings	940 Mbyte optical discs	45	9,000 hours
Questionnaires, Subjective Ratings	Sheets of paper	966	483 driver-days

At the time of writing this report, it was planned to put the database into an archiving system and making it available to researchers. However, methods and details have yet to be established.

Table 7. The reduced data set used for analysis

SOURCE	REDUCTION METHOD	RESULT	SIZE
Video	1 minute of 30	Paradox™ DBMS file	14,697 records of 50 fields
Truck computer	Edit	Flagged SYSTAT™ file	560 Mbytes on 1 optical disk
Truck computer	1 minute epoch statistics	Paradox™ DBMS file	335 Mbytes divided into 335,000 records
Physiology: sleep	Manual R-K scoring	Sleep hypnogram	14 Mbytes of ASCII text
Physiology: driving	Spectral analysis scoring	Paradox™ DBMS file	Critical events
Questionnaires, Ratings	Keyed data to disk	Paradox™ DBMS file	4 Mbytes

SECTION 4. RESULTS, CONCLUSIONS, AND RECOMMENDATIONS

This section begins with a synopsis of the major study results concerning the following: driver time off; sleep, naps, and physiological measures; driver drowsiness judged from face video recordings; driver drowsiness judged from physiological measures; driver performance and vehicle control measures; driver performance on surrogate performance tests; relative importance of fatigue factors; research methods; and the data base.

From these results, a set of general conclusions is drawn and a related discussion is provided concerning their relationships to several key issues surrounding CMV driver fatigue: time of day (circadian) effects, duration of driving time, cumulative fatigue across days, daily off-duty requirements, quality and quantity of sleep obtained, the four driving schedules and observed differences in fatigue, drowsiness during driving, napping, effect of mid-trip breaks, driver self-awareness of fatigue, individual differences, sleep apnea, age and fatigue, countermeasures and usefulness and limitations of fitness-for-duty testing found within the context of this study.

The chapter concludes with a set of recommendations for fatigue countermeasures which follows the taxonomy developed in the literature review: alertness maintainers controlled by the driver or externally, and alarms based on driving performance or on driver physiological state.

The driving schedules used as data collection Conditions are presented in Table 2 and should be referred to here as a refresher to facilitate understanding of this section.

SYNOPSIS OF MAJOR STUDY RESULTS

DRIVER TIME OFF

Much of the variability in total time off (difference between trip start time and the preceding trip's end time) in all four Conditions was accounted for by activities that occurred after the drive, probably due mainly to time required for commuting to the sleep site, eating and socializing. Some drivers chose to spend substantial time socializing after work, while others emphasized starting their principal sleep period.

The primary between-Condition differences in total and post-drive time off were the result of the differences between the work/rest schedule in Condition C1-10day and the other schedules, as planned.

The drivers in Conditions C2-10rotating, C3-13nightstart and C4-13daystart had available, on average, approximately 8.6 to 8.9 hours of continuous off-duty time (excluding time required for study protocol related demands), while those in Condition C1-10day had approximately 10.7 hours. Case-by-case examination of the sleep laboratory and truck terminal technicians' notes might permit detailed quantitative estimates of the various factors accounting for off-duty time for all Conditions.

The study design was developed to comply with existing hours-of-service regulations. The drivers, all of whom participated voluntarily, performed tasks related to the study protocol that required a nominal amount of time during their off-duty periods. If this time is considered to be "on duty," then, in 88% of the cases, drivers had at least the minimum off-duty time required. If this time is considered to be "off-duty," then, in 97% of the cases, drivers had at least the minimum off-duty time required.

SLEEP, NAPS, AND PHYSIOLOGICAL MEASURES

The driver fatigue study used standard sleep laboratory protocols. These protocols did not appear to influence sleep physiology and did not appear to significantly impact the quality and quantity of sleep obtained based on post-sleep questionnaire data obtained from the drivers and other polysomnographic databases available in the literature.

The average difference over all Conditions between reported "ideal" sleep time and observed time in bed was 1.97 hours. (In a "sleep history questionnaire," drivers were asked to complete the statement "My ideal amount of sleep is _____ hours.")

Drivers in Condition C3-13nightstart spent 262 minutes (4.37 hours) in bed per principal sleep period, which was significantly shorter (40-85 minutes) than the time in bed for Conditions C1-10day, C2-10rotating, and C4-13daystart. The schedule in Condition C3-13nightstart required drivers to depart on their runs in late evening and arrive back at the start terminal in the afternoon.

Polysomnography of principal sleep periods documented low sleep latencies (i.e., drivers fell asleep quickly), low wake times after sleep onset (i.e., drivers were awake for only brief periods of time after they first fell asleep) and high sleep efficiencies (i.e., drivers slept most of the time they were in bed). These were all consistent with the observed amount of time spent in bed and the amount of sleep that the drivers obtained. In addition, the amounts of wake time after sleep onset were quite low relative to values that are considered in the normal range and were consistent with reduced time in bed and time asleep. The drivers' sleep, with the exception of low total sleep time, was unremarkable (i.e., normal) in terms of sleep physiology and structure.

Employing standard clinical polysomnographic, oximetric and pulmonary ventilatory measurement methodologies, obstructive sleep apnea was detected in 2 of the 80 drivers (2.5%) participating in the study. Sleep apnea is a sleep disorder where breathing difficulties disturb sleep and therefore can cause sleep debt and work period drowsiness. These two drivers, both of whom were in the older age group of Condition C1-10day, had apnea rates in the range of 10 - 30 per hour of sleep.

There appeared to be no strong rationale for excluding the polysomnographic data of the two drivers with sleep apnea from the sleep analysis. Their total sleep time and sleep architecture variables were not grossly different from the other older drivers of Condition C1-10day. Thus, they were included in the analysis.

The averaged Stanford Sleepiness Scale (SSS) ratings showed a clear rise and fall in reported sleepiness that recurred every 24 hours. There were no easily interpretable or systematic differences in average SSS ratings across Conditions and age groups.

Of the 80 drivers, 35 (44%) took naps. Drivers who did elect to nap augmented their sleep obtained in principal sleep periods by about 27 minutes, which amounted to an 11% increase in their average sleep time.

Episodes judged drowsy in video records were often precursors to drivers deciding to take naps. Thus it appeared that this behavior was replacement or compensatory napping, taken in response to self-perceived sleepiness.

DRIVER DROWSINESS JUDGED FROM FACE VIDEO RECORDINGS

The continuous video recordings of the driver's face were sampled every half hour of trip time, and various observations of the road and driver were made, including whether the driver appeared to be drowsy based on eyelid closures and other criteria. When drowsiness was noted, the video was viewed from 30 minutes before to 30 minutes after the initially detected episode, for a total of ten 6-minute samples. The sampling and analysis procedure permitted estimating prevalence and duration of drowsiness while driving during this data collection. Prevalence of drowsiness refers to the proportion of face video samples judged drowsy in a given interval of time.

> There were 1,989 6-minute periods of driving that were judged drowsy (198.9 hours drowsy/4,080 hours total = 4.9%, a lower bound for prevalence of drowsiness averaged over the entire data collection). Of these, 1,646 periods (82%) occurred during the hours 1900 - 0659.
>
> The observed prevalence of drowsiness averaged over all four Conditions by hour of the day formed a distinct peak about eight hours wide spanning late evening until dawn. This finding is consistent with the findings of many studies of the effects of circadian rhythms.
>
> The frequency of drowsiness run lengths (e.g., consecutive drowsiness episodes) could be well described by the exponential density function ($R^2 = 0.92$). Longer runs of drowsiness occurred with decreasing frequency. The minimum run length was 1, the mode (i.e., the most frequent) was 1, the median run was four 6-minute drowsy epochs (=24 minutes), the mean was 6.44, and the maximum run length was thirty-seven 6-minute drowsy epochs (3.7 hours).
>
> About two-thirds of the drivers were judged drowsy in at least one video sample; however, 11 drivers (14%) accounted for more than half (54%) of all the observed drowsy samples. In other words, within the confines of the study some drivers displayed many more drowsy episodes than others. Because there was no long-term observation of subjects, it cannot be discerned whether these observed individual differences were reflective of driver traits (i.e., long-term, stable individual differences in physiology and/or performance) or of driver states (short-term differences related to recent sleep or other transient events). Of course, both traits and states may be operative. Future research should address the trait versus state issue; findings in support of the trait hypothesis would imply, for example, that selection tests might be developed to identify drivers most resistant to fatigue.
>
> There was no difference in the prevalence of drowsiness observed in the video data during comparable (i.e., "daytime") trip segments of the 10-hour and 13-hour trips. Thus the longer

working and driving periods associated with the 13-hour route, which might be expected to produce adverse effects, did not do so with respect to daytime drowsiness.

There was greater prevalence of drowsiness in Condition C2-10rotating, trips 4 and 5 (compared to trips 1, 2, and 3), probably because the rotating schedule had caused these last trips, on average, to be driven through the night. Although disruption of circadian rhythms and cumulative fatigue probably contributed, the major factor seemed to be time of day.

The prevalence of drowsiness in Conditions C3-13nightstart and C4-13daystart was markedly greater during night driving. It was concluded that this was due mostly to the low period in the drivers' circadian rhythms, although the effect was probably amplified by sleep debt. This relative night versus day effect was not examined for Conditions C1-10day and C2-10rotating, the former because it did not include night driving and the latter because it included a limited amount of night driving.

There were about 600 times as many minutes of face video-indicated drowsiness than indicated by PSG methods. Judgments of face video apparently were far more sensitive for detecting drowsiness while driving than methods based on PSG (polysomnographic) measurements. The PSG measurement may have revealed a worse (by comparison with the face-video judgments) and infrequently-occurring condition referred to in this study as "PSG-Drowsy Driving." However, these results may be reflective of the relative sensitivities of the two methods in detecting drowsy driving.

DRIVER DROWSINESS JUDGED FROM PHYSIOLOGICAL MEASURES

The PSG (polysomnographic) data showed that there were two trips, involving different drivers -- an incidence of about 0.6% of observed trips and about 2.5% of observed drivers -- which contained a number of intermittent episodes during which brain electrical activity indicated a state identified in this study as PSG-Drowsy Driving. Neither driver had sleep apnea. These periods amounted to just over 19 minutes out of 244,667 minutes (nearly 4,000 hours) of driving PSG analyzed, or 0.008% of that time.

For one driver in Condition C3-13nightstart (Case 1), the drowsiness occurred during an initial period of 60 seconds and a second period of 20 seconds. These two periods began approximately 2 hours and 3 minutes and 4 hours and 18 minutes, respectively, after the start of the last outbound trip of the week, at 0224 and 0438. There were indications of acute or cumulative fatigue from the driver's preceding and subsequent sleep periods. The video of the driver's face showed that the driver was mildly drowsy during these two episodes. This driver was in his 20s. The times at which the episodes occurred were at the circadian low point of the drivers in this Condition.

For one driver in Condition C4-13daystart (Case 2), the drowsiness occurred intermittently over a period of 29 minutes. This period began approximately 10 hours and 15 minutes after the start of the first outbound trip of the week, at about 2323. There was no indication of acute or cumulative fatigue from the driver's preceding and subsequent sleep periods. The video of the driver's face showed that the driver was drowsy during these episodes, and that in two instances, the truck drifted across a lane boundary after the driver’s eyes had been shut for six or more seconds. This driver was in his 30s. The times at which the episodes occurred were at the circadian low point of the drivers in this Condition.

The Case 1 episodes of PSG-Drowsy Driving (which happened at the start of the last shift of the week, in Condition C3-13nightstart) occurred in a manner that could be explained by acute or cumulative fatigue, which was exhibited in this driver's preceding and subsequent sleep periods, combined with the low in circadian rhythm at the times during which the episodes occurred.

The Case 2 episodes of PSG-Drowsy Driving (which happened toward the end of the first shift of the week, in Condition C4-13daystart) did not occur in a manner that could be explained by acute or cumulative fatigue. The times at which the episodes occurred were, however, during the low point in circadian rhythm. This, possibly combined with time on task effects and/or some other unknown factor, could explain the occurrence.

DRIVER PERFORMANCE AND VEHICLE CONTROL MEASURES

Steering wheel variability was adjusted for route-dependent effects by subtracting the all-driver average steering variability associated with each mile of road, thus reducing the variation associated with road curvature.

Lane position was established using a lane tracker which measured the distance between the right-hand side of the vehicle and the right-hand lane line.

Steering wheel variability was strongly and reliably affected by location on the route. Steering wheel variability therefore must be corrected for route-dependent effects if it is to achieve its full potential as an indicator of driver fatigue.

When adjusted for route-dependent effects, steering wheel variability for video drowsy epochs in Conditions C3-13nightstart and C4-13daystart proved to be greater than the same measure for not-drowsy epochs. Similar analyses were not conducted for Conditions

C1-10day and C2-10rotating because of the limited amount of drowsiness seen in those Conditions. Steering power spectral density and steering wheel reversals should also be evaluated to further assess these measures as indicators of fatigue and loss of alertness.

Lane tracking standard deviation readings were systematically greater in the 13-hour Conditions by comparison with the 10-hour Conditions. The reasons for this are not completely clear because of confounding factors associated with different routes and vehicles. Additional study of the existing data, especially video, will be required to fully resolve this issue.

Drowsiness (as observed in face video records) was associated with increased lane tracking variability in all four Conditions, but was statistically significant only in Condition C1-10day and Condition C3-13nightstart. Drowsiness had been hypothesized to cause increased lane tracking variability, and it was directly observed to do so in several video recordings of driving. Therefore, it was concluded that lane tracking variability is a valid indicator of driving performance, and that drowsiness causes increased lane tracking variability.

DRIVER PERFORMANCE ON SURROGATE PERFORMANCE TESTS

The three tests included in the surrogate performance test battery provide four measures (CS, RVS, LAPSES and CTT) as follows: (1) Code Substitution (CS) Test - the number of correct responses per second; (2) Simple Response Vigilance Test (SRVT) - median time to respond to a random stimulus, from which are derived the two measures (a) Response Vigilance Score (RVS), the reciprocal of median response latency of all responses, and (b) LAPSES, the number of response latencies greater than 500 milliseconds; Critical Tracking Test (CTT) - the median score for the critical value of instability (at which the driver is unable to compensate for the displayed error) of the five CTT trials per test session. The large amount of data obtained from the performance tests allowed statistically reliable estimates of small effects.

In the 10-hour Conditions (C1-10day and C2-10rotating), there were five trips and four daily administrations of the performance test battery. There were four trips and three daily administrations of the performance test battery in the 13-hour Conditions (C3-13nightstart and C4-13daystart). The number of trips were limited by the hours-of-service rules in the respective jurisdictions. There were only three performance test administrations per trip in the 13-hour Conditions, because of schedule constraints imposed by the greater distance driven and hours-of-service regulations.

Code Substitution (CS) scores improved with practice, were higher in younger drivers, and were better during the day for all Conditions.

Drivers improved their CS scores with each successive trip in all four observational Conditions because of practice effects.

Despite the practice effects, CS performance degraded during trips in Conditions C1-10day , C2-10rotating, and C4-13daystart. In Condition C3-13nightstart, however, performance was better at trip end, since the earlier administrations were performed during drivers' circadian low. In Condition C3-13nightstart, circadian effects outweighed time-on-task effects. The Code Substitution Test appears to be sensitive to behavior changes commonly associated with acute fatigue and loss of alertness.

The 10-hour drivers did better on the CS test than the 13-hour drivers. They also showed greater improvement across the first three trips. This result may have been due to greater fatigue in the 13-hour Conditions.

There were RVS score decrements from first to last trips of the week in all four observational Conditions. In general, RVS scores were lower at the end of trips than they were at the start in Conditions C2-10rotating, C3-13nightstart , and C4-13daystart. Lower scores were the result of increased response latencies associated with acute fatigue. RVS scores were sensitive to both acute (within trips) and cumulative (across trips) fatigue, despite the tendency toward poor performance in daylight, an uncontrolled nuisance variance, and despite extensive missing data.

RVS scores of the 13-hour observational Conditions were better than those of the 10-hour Conditions. The results were confounded by varying light levels between the 10-hour and 13-hour Conditions. In spite of dilution by nuisance variance associated with uncontrolled ambient light levels, this simple metric was sensitive to performance losses both during and across trips.

The number of LAPSES increased from first to last trip in the two 10-hour Conditions. In the 10-hour rotating-start Condition, drivers generally experienced more Lapses at the end of trips, than they did at the start. There were no statistically significant differences in number of LAPSES from trip-start to trip-end in the 13-hour Conditions. The lack of meaningful changes in LAPSE data may have resulted from the uncontrolled influence of time-of-day on stimulus intensity (e.g., CRT contrast ratios).

There were more LAPSES in the 10-hour Conditions than in the 13-hour Conditions. Performance may have been influenced by a greater amount of sunlight in the 10-hour Conditions than in the 13-hour Conditions, with increased glare on the test's CRT display.

This glare was an artifact which was due to uncontrolled variation in ambient light levels during data collection.

For both the 10-hour and 13-hour Conditions, CTT scores did not vary by Condition, trip or administration.

The Simple Response Vigilance Test (SRVT), a very simple test from which can be derived RVS scores and number of LAPSES, may be the best performance test index of cumulative fatigue. It proved sensitive to vigilance and perceptual-motor speed changes that typify loss of alertness and fatigue.

Concerning performance recovery after the 54-minute mid-trip break at the turnaround terminal in the 10-hour Conditions, only Code Substitution (CS) performance improved. The other performance test scores failed to show statistically significant recovery. Recovery after the mid-trip break could not be estimated in the 13-hour Conditions, since the test battery was not administered at the end of the break in those Conditions.

For all Conditions, there was little correlation between driver self-ratings on the Stanford Sleepiness Scale and performance test scores.

RELATIVE IMPORTANCE OF FATIGUE FACTORS

The relative impact of such factors as hours of driving, days of work, and time of day on measures like drowsiness, lane tracking variability, surrogate test performance, and amount of sleep was established with a non-parametric measure of association, which requires only that measurement on an ordinal scale be achieved. That is, measured values are only required to be placed in rank order according to hypothesis, without need to specify exactly where along a scale of measurement they lie. This provides a robust method of testing according to their strength of association (e.g., rank order correlation) with criterion measures (e.g., selected performance measures such as video drowsiness, lane tracking variability, etc.).

A mathematical model was developed of the rhythmic or circadian effects seen in the drowsiness data of Conditions C3-13nightstart and C4-13daystart based on a two-peak, 24-hour pattern discussed by Mitler (1990) in relation to sleep, mortality, and human error. Parameter values were calculated to optimize the fit of the model to the data from this study, and the goodness of fit was evaluated. This model provided another means of testing the relative importance of time of day, time on task, and cumulative trips with respect to observed drowsiness.

Night driving (0000-0600) was associated with worse performance on each of four important criterion variables (drowsiness, lane tracking, code substitution, and sleep length), whereas hours of driving and number of consecutive trips had little or no relationship to those criterion variables. It was concluded that time of day was a far better predictor of decreased driving performance than time on task or cumulative number of trips.

Drivers' self-ratings differed sharply from the other criterion variables (drowsiness, lane tracking, code substitution, and sleep length). Unlike the other criteria, drivers' self-ratings correlated significantly with trip segments ranked primarily by hours of driving, and correlated significantly with trip segments ranked primarily by cumulative number of trips, but the self-ratings did not correlate significantly with trip segments ranked by percent of night driving. Thus, the self-ratings were not very good indicators of drowsiness, but they may have been indicative of increasing stress or compensatory effort that signaled fatigue or loss of alertness.

The developed mathematical model based on a two-peak, 24-hour, circadian pattern had a strong relationship to observed prevalence of drowsiness in Condition C3-13nightstart and Condition C4-13daystart judged from face video recordings. It was concluded that this was because of the strong time of day (e.g., circadian) influence on observed prevalence of drowsiness. The model is a useful descriptor of drowsiness observed in these data because it provides smooth, continuous estimates for any time of day. A model of this nature could be used as a training aid and work shift planning aid.

Neither elapsed time since trip start nor cumulative number of trips contributed significantly to the accuracy of the rhythmic (circadian effects) model's estimates of drowsiness observed in Condition C3-13nightstart and Condition C4-13daystart. It was concluded that the predominant factor relating to observed drowsiness was time of day.

"Elapsed time since start of trip" was considered important because of its strong relationship to observed driver self-reports, even though those self-reports differed sharply from the objective measures. We concluded that these self-reports may have indicated increasing stress or compensatory effort that signals fatigue or loss of alertness, and that the drivers had diminished motivation and abilities to remain alert by the end of their trips (Brown, 1994; Dinges et al., 1994).

RESEARCH METHODS AND THE DATABASE

It was possible to conduct a field study with substantial numbers of commercial drivers (80), hauling revenue freight on many trips (360), employing instrumentation to record several aspects of vehicle control, surrogate test performance, driver physiology, face and road video, and sleep studies for each principal sleep period. This was accomplished without any motor vehicle or other accidents.

Concurrent measurement of many variables was very important for data interpretation. It allowed results to be derived that could not be obtained with confidence from analysis of one or two measures.

The database created in this study will be useful for many years, to address questions and/or to employ analysis methods beyond the scope of the present study.

DISCUSSION BY MAJOR CMV DRIVER FATIGUE ISSUES

Project results and conclusions are reviewed and discussed below as they relate to major CMV driver fatigue issues.

TIME-OF-DAY (CIRCADIAN) EFFECTS

The strongest and most consistent factor influencing driver fatigue and alertness in this study was time of day because of the effects of 24-hour biological rhythms, known as circadian rhythms. Drowsiness, as observed in face videos, was markedly greater during night driving than during daytime driving. Peak drowsiness occurred during the eight hours from late evening until dawn. Night driving (e.g., from midnight to dawn) was also associated with worse performance on four important criteria: proportion of video-drowsy analysis periods; average lane tracking standard deviation; incremental differences in Code Substitution test scores between the outbound and inbound segments of a trip; and average physiologically measured total sleep obtained during the principal sleep period prior to a trip. Time of day was a much better predictor of decreased driving performance than hours of driving (time-on-task) or the cumulative number of trips made.

The drivers' Code Substitution (CS) test scores were affected by circadian rhythms, as were their self-ratings of sleepiness -- both were worse at night. However, the most pronounced effect of circadian rhythms was on the number of video recordings of drivers who were judged drowsy while driving. During driving periods between 2200 and 0600 hours, there was an eightfold increase, compared to daytime levels, in the amount of video samples judged drowsy. There were a total of 1,989 6-minute periods of driving (198.9 hours out of 4,080 hours of video, or 4.9 percent, a lower bound for prevalence of drowsiness) that were judged drowsy. Of these, 1,646 periods (82%) occurred during the hours 1900 - 0659. Interestingly, the NTSB (1995) study of single-vehicle truck accidents that were likely to have been fatigue related (limited further to those in which the driver survived and in which the previous 96 hours could be reconstructed) found a total of 62 out of 107 accidents that were fatigue-related, of which 52 (83.9%) occurred at night (i.e., between 2200 and

0800) and 10 occurred by day, a ratio of 5.2:1 for the observed number of night versus day fatigue-related accidents.

The degree of drowsiness was not judged on any scale, but it is believed that these drivers ranged from very mildly drowsy to extremely drowsy. The question naturally arises, what is meant by "extremely drowsy"? It is difficult enough to measure effects of fatigue at all, but a critical weakness of most driver fatigue research is the difficulty of answering the question, was the fatigue measured in a given situation excessive? Normative standards or "pass/fail" scores are usually lacking for measures such as steering wheel and lane tracking variability, code substitution and simple response vigilance test scores. However, because driver and forward-looking video recording was used in this study, it was possible to conclude, from visual observation of the drivers' eyes and of lane position, that some drivers were extremely drowsy and would have been well advised to park in a safe location and nap. Unfortunately, project resources did not permit a detailed case-by-case study of naps taken subsequent to those periods to assess the proportion of drivers who took this remedial measure.

Almost all of these periods of extreme drowsiness occurred at night, when drivers' circadian rhythms made them most susceptible. Their tendency to drowsiness was probably made worse by the low amount of sleep, which involved a number of factors (discussed elsewhere), not the least of which was the necessity of the night drivers to take their principal sleep periods during the day, also contrary to their natural body rhythms. Thus the circadian rhythm acts against the night driver both while working and while trying to sleep. In addition, there are fewer alerting environmental stimuli at night due to darkness and the presence of fewer vehicles on the roadways.

DURATION OF DRIVING

There was no difference in the prevalence of drowsiness observed in video records of comparable daytime segments of the 10-hour and the 13-hour trips. Lane tracking performance was better in the 10-hour than the 13-hour Conditions. However, this difference may have been due to greater fatigue, to differences between the routes and/or the vehicles, or a combination of these. In the surrogate tests, cognitive performance (via Code Substitution) was better in the 10-hour Conditions, vigilance and reaction time (derived from the Simple Response Vigilance Test) were better in the 13-hour Conditions (probably because of loss of display contrast associated with greater amounts of sunlight in the 10-hour Conditions), and hand-eye coordination (from the Critical Tracking Task) did not show Condition-related variation. Self-ratings of fatigue level on the Stanford Sleepiness Scale correlated positively with time-on-task, indicating that drivers may have the subjective feeling of increasing fatigue with increasing time-on-task even if there are no strong performance changes. There was little correlation between Stanford Sleepiness Scale self-ratings and performance test scores.

The hypothesis that driving performance declines with time-on-task is the cornerstone of the present hours-of-service regulations, first introduced in 1937. This hypothesis has dominated driver fatigue research for decades. In this study, an analysis of the relative importance of factors was conducted, and where drowsiness or sleep tendency is concerned, time of day was shown to be far more important than time-on-task or cumulative number of trips.

An awareness of the importance of sleep and of natural body rhythms has grown apace with the rapid scientific and medical developments in these areas. However, care must be taken not to let the pendulum swing too far the other way, leading to dismissal of time-on-task as irrelevant. The data collected in this study showed a strong relationship between "elapsed time from trip start" and self-ratings. These changes of self-ratings with time-on-task may reflect increasing fatigue, where fatigue is defined as a subjective experience distinguishable from physiological impairment, as Brown (1994) concluded. They may indicate compensatory effort (Dinges et al., 1994). According to these views, worsening self-ratings would not necessarily be correlated with objective performance measures, but may nonetheless indicate a decrease of reserves and increase in risk.

CUMULATIVE FATIGUE ACROSS DAYS

There was some evidence of cumulative fatigue across days of driving. For example, performance on the Simple Response Vigilance Test declined during the last days of all four Conditions. Also, drivers tended to rate themselves as more fatigued across multiple trips. However, cumulative number of trips was neither a strong nor consistent predictor of fatigue across different measures. Although more apparent drowsiness was noted in video recordings made in the last two trips of Condition C2-10rotating, those trips were, on the average, driven at night and the more apparent drowsiness may be a reflection of circadian effects. The Stanford Sleepiness Scale self-ratings of sleepiness increased as drivers progressed through successive trips within Condition C2-10rotating, but the trends were unclear in Condition C3-13nightstart and Condition C4-13daystart.

DRIVER TIME OFF

It was seen that much of the variability in total time off in all four Conditions was accounted for by activities that occurred after the drive. The reason for the strong relationship between post-drive time off and total time off is simply a mathematical artifact of normal trip scheduling as follows.

The time off between duty periods is, fundamentally, made up of three segments: end of duty until bedtime, time in bed, and time of arising until duty start. The time that the duty period starts is the least flexible point in the schedule. Usually, management has fixed that time as a departure time and planned other activities around it. Changing that time is usually not an attractive prospect. Thus, the driver, in turn, plans his or her time in bed based upon the scheduled duty start time. The time in bed is somewhat more flexible than the duty start time. However, with tight schedules such as those used here in Conditions C2-10rotating, C3-13nightstart and C4-13daystart, there is not much flexibility if adequate or nearly adequate sleep is to be obtained or at least attempted.

Finally, there is some flexibility forced into the length of the post-duty period by factors within and outside the control of the driver. The factors within the driver’s control were mentioned above. Those outside the driver's control are all of those that affect the length of the previous trip: traffic; weather; interactions with shippers, and receivers and the driver's own management; inspections; breakdowns; etc. If the duty period ends earlier than expected, both the post-trip time off and the total time off are longer than expected. If the duty period ends later than expected, both the post-trip time off and the total time off are shorter than expected. This normal covariation is the most likely source of the significant statistical correlation we observed between post-drive time off and total time off.

There may or may not be a fatigue countermeasure lesson here for use in driver education. The lesson is fairly obvious and may have been internalized by most drivers already. The rationale of the lesson follows. The driver cannot predict when duty periods will be longer than expected. In a tight schedule, the driver should be as rested as possible at the start of each duty period. That state will make the driver as resilient as practical to fatigue-inducing long duty periods. Thus, the driver should take advantage of short duty periods and subsequent, longer-than-expected, rest periods to obtain a normal night of sleep or to allow for an even longer sleep. Fatigued, sleep-deprived people will sleep longer than their normal amount of time. The phenomenon is labeled "recovery sleep" by the sleep research community. This argument is not for "storing up sleep." Rather, the argument is for the driver to allow time for normal or for recovery sleep when a duty period ends earlier than expected.

There may also be a fatigue countermeasure lesson here for trucking schedulers. In United States Air Force (USAF) transport flight operations in the field, an aircraft commander has the responsibility and the authority to lengthen crew rest for his or her flightcrew if the period of time off between flight-duty periods does not allow adequate time for the sleep period specified by regulations. The effect of an aircraft commander's decision to extend a crew rest period is to delay the scheduled departure of the aircraft. Such decisions are not popular, but they are accepted within the USAF management community as being essential for safe flight. Thus, the scheduling system is structured to deal with these perturbations.

Similarly, truck dispatchers must deal daily with major perturbations in planned pickups and deliveries. The lesson here may be to incorporate driver sleep as one of the expected perturbations with which the scheduler must deal. If a driver informs the scheduler that he or she has not obtained adequate sleep due to a longer than expected duty period or by duty-period delays, then the dispatcher could use that information to adjust the driver's subsequent schedule.

The key to this concept is the length of the sleep period, not the length of the off-duty period, between duty periods. The drivers in the study reported an ideal sleep time of 7.2 hours; when other activities undertaken during off-duty periods within or outside of a study setting are taken into account (eating meals, taking care of personal needs, commuting), eight hours off between duty periods might not provide enough time to obtain adequate sleep. This is especially true when circadian rhythms are added as a consideration. If the driver cannot get to sleep because it is the middle of the day, then more than eight hours off duty might well be needed for the driver to obtain adequate sleep between duty periods.

DAILY SLEEP PERIODS

Overall, drivers obtained about 2 hours less time in bed and 2 ½ hours less actual sleep than their reported ideal daily sleep requirement. However, the shortfall was greatest for the Condition C3-13nightstart drivers, who had a shortfall of 2.3 hours relative to time in bed and 3.4 hours relative to actual sleep. Drivers' average reported ideal sleep time was 7.2 hours per principal sleep period; the average observed time in bed over the course of the study was 5.2 hours and time asleep was 4.8 hours. For each of the four Conditions, the approximate amounts, on average, of continuous off-duty time (excluding time required for the study protocol), time in bed, and clinically measured sleep time were as follows:

Condition C1-10day: 10.7 hours off-duty, 5.8 hours in bed, 5.4 hours asleep.
Condition C2-10rotating: 8.7 hours off-duty, 5.1 hours in bed, 4.8 hours asleep.
Condition C3-13nightstart: 8.6 hours off-duty, 4.4 hours in bed, 3.8 hours asleep.
Condition C4-13daystart: 8.9 hours off-duty, 5.5 hours in bed, 5.1 hours asleep.

The observed sleep shortfall (compared with the driver reported ideal) could have been due, in part, to a reduction of free time by requirements of the study protocol. The study setting also created an opportunity for socializing with other drivers that might not exist in normal driving. In addition, the drivers often did not organize their off-duty time to obtain the maximum possible sleep. Further, although drivers reported what they considered to be their ideal sleep time, they were not asked and it is not known whether they usually obtained their stated ideal sleep.

Time-in-bed was lowest for the three Conditions (C2-10rotating, C3-13nightstart, and C4-13daystart) that permitted the least amount of continuous off-duty time (about 8.6 to 8.9 hours excluding time required for the study protocol). Nevertheless, even in Condition C1-10day, which permitted about 10.7 hours off-duty between trips (excluding time required for the study protocol), the average time-in-bed and time asleep were only 5.8 and 5.4 hours, respectively.

The lower ratio of sleep time to time in bed for Condition C3-13nightstart probably reflects circadian disruptions of sleep pattern in comparison with the other Conditions. This Condition was the only one that consistently required drivers to sleep during the daytime.

QUANTITY AND QUALITY OF SLEEP OBTAINED

The quantity of sleep obtained by the subjects in their principal sleep periods was low. As noted earlier, drivers obtained an average of about 2 hours less time in bed and 2 ½ hours less actual sleep than their reported daily "ideal." The average time-in-bed was 5.2 hours (average sleep time of 4.8 hours) versus their reported ideal amount of 7.2 hours of sleep. The shortest average time-in-bed (4.4 hours) and actual sleep time (3.8 hours) were associated with Condition C3-13nightstart.

Sleep quality was high. All of the drivers obtained efficient, normally structured sleep as judged by formal clinical criteria. Of the 5.2 average hours in bed, the drivers were actually asleep for an average of 4.8 hours. The average sleep efficiency (sleep time/time-in-bed) was 0.92; levels above 0.90 are often observed in people who have no trouble sleeping and in people who are sleep deprived. The average amount of time awake after sleep commenced was 25 minutes; this value is considered low relative to values in the normal range (less than 60 minutes for adult men) and is also consistent with reduced time in bed and with sleep deprivation.

The study design was developed to comply with existing U.S. and Canadian hours-of-service regulations. It was expected that the drivers would get adequate sleep. They did not. The reasons for this are not simple or clear. The data collection protocol nominally took about 0.9 hour total from the drivers' post-work-period time off (i.e., time between trip end and trip start), leaving the Condition C2-10rotating, C3-13nightstart, and C4-13daystart drivers with about 8.7 hours on average. Some of that time was spent on the necessary activities of commuting to and from work, eating, and personal hygiene. Several instances were observed of social and recreational activities (e.g., talking, reading, watching television) that seemed to be time ill-spent in an intense work schedule where sleep should have had higher priority. In any event, although their minimum requirement for sleep could not be established precisely, these drivers certainly got less sleep than they needed as judged by formal clinical criteria (overall average was 4.8 hours per principal sleep period).

These drivers were observed on a nominal 21-hour work-rest cycle (Condition C2-10rotating) that caused them to drive between 2200 and 0600 during their last few trips, and on nominal 24-hour work-rest cycles (Conditions C3-13nightstart and Condition C4-13daystart) that involved driving between 2200 and 0600 during most trips. During this night driving, there was on average an eightfold increase, compared to daytime levels, in the amount of driving judged "drowsy" from the face video recordings.

The prevalence of drowsiness observed would likely have been less if the drivers had obtained adequate sleep. A statistically significant negative correlation was observed between sleep length and drowsiness, but there was no observational Condition where all drivers obtained adequate sleep and performed night driving, so there are insufficient data from this study to estimate "normal" levels of nighttime drowsiness. Likewise, there are insufficient data to estimate reliably what the effect would have been if the protocol had not claimed an hour of the drivers' off-duty time.

DROWSINESS DURING DRIVING

Video ratings were much more sensitive for detecting drowsiness while driving than were polysomnographic (PSG) measures. Because project resources did not permit review of 4,000 hours of video tape, the video recordings were systematically sampled at 30-minute intervals and drowsy episodes were judged in 6-minute periods from 30 minutes before to 30 minutes after their occurrence. Approximately 4.9% of the sampled face video segments were scored as drowsy based on trained reviewers' assessment of such factors as eye movement, eyelid position, yawns, stretches, and startles. The proportions of video data scored drowsy were much greater at night than during the day or evening.

All EEG and EOG data were manually analyzed by trained polysomnographic technologists. PSG analysis indicated that there were two trips, involving different drivers (an incidence of about 0.6% of observed trips and about 2.5% of observed drivers), that included a number of intermittent episodes that were identified in this study as PSG-Drowsy Driving. These periods amounted to just over 19 minutes out of the 244,667 minutes of driving analyzed (0.008%). During these periods, the drivers' exhibited EEG and EOG patterns that would have been consistent with clinical (Rechtschaffen and Kales, 1968) criteria for Stage 1 sleep (the initial, shallowest, sleep stage) if the drivers had been in bed in a dark room.

Both episodes of PSG-Drowsy Driving were also associated with the drivers appearing drowsy on the corresponding video records. The EEG measurement may have revealed a worse (by comparison with the face-video judgments) and infrequently-occurring condition. However, these differences may be more reflective of the relative sensitivities of the two methods of detecting drowsy driving. Some sleep researchers consider Stage 1 to be a drowsy or marginal condition between wakefulness and sleep. Other sleep researchers consider Stage 1 to be actual sleep with all that the term sleep implies in terms of sensory and behavioral responsiveness. Based on the current state of knowledge concerning human responsiveness, it is justifiable to be concerned about the safety aspects of driving a motor vehicle while the cerebral cortex appears to be in a stage of sleep. However, a driver may still show overt behaviors that indicate drowsiness, and be in a state of reduced alertness and/or awareness in the absence of an EEG pattern that definitively indicates the onset of sleep.

A comparison of steering and lane tracking performance for video-rated drowsy versus non-drowsy epochs indicated that drowsiness was associated with more erratic steering (greater steering wheel angle variability) and poorer lane tracking (increased standard deviation of lane position), both of which have obvious implications for driving safety.

Not surprisingly, there was a negative correlation between the length of the principal sleep period and the amount of drowsiness during the next driving trip ($r = -0.37$, $p < 0.05$). However, it was not possible to estimate the "normal" level of drowsiness during driving since there were no Conditions where all drivers obtained adequate sleep.

Although there were video, PSG, and driving performance indications of driver drowsiness, there were no instances of crashes in the study.

NAPPING

Of the 80 drivers, 35 (44%) took at least one nap during a duty cycle that contained clinically scoreable sleep. Drivers who elected to nap increased their sleep obtained in principal sleep periods by an average of 27 minutes, which amounted to an 11% increase in average daily sleep time. Drowsiness, as evident in face video recordings, was often a precursor to the driver deciding to take a nap. Thus it appeared that this behavior was replacement or compensatory napping, taken in response to self-perceived sleepiness.

Because 45 of the drivers did not nap, and there were only 63 naps taken over the 360 trips in the study, no analyses were performed to determine whether these driver naps resulted in post-nap improvement in alertness and performance. This is one of many important questions which might be addressed by future detailed case study analysis of the data collected in the Driver Fatigue and Alertness Study (DFAS).

EFFECTS OF MID-TRIP BREAKS

In the 10-hour Conditions (C1-10day and C2-10rotating), drivers self-administered the surrogate performance tests both at the beginning and the end of their mid-trip turnaround break. The only test demonstrating improved post-break performance was the Code Substitution test. The other performance tests failed to show a statistically significant recovery effect.

DRIVER SELF-AWARENESS OF FATIGUE

There was little correlation between driver subjective self-ratings of alertness/sleepiness and concurrent performance measures. It appears that drivers are not very good at assessing their own levels of alertness; there was a tendency for drivers to rate themselves as more alert than the performance tests indicated.

On the other hand, there was a positive correlation between self-ratings of fatigue and both the number of hours of driving within a trip and the cumulative number of trips made. Perhaps these factors affected the subjective experience of fatigue, reflecting increasing stress or compensatory effort rather than objective performance. Or, perhaps drivers were basing their self-ratings in part on a logical expectation that these factors would increase fatigue and they would thus be led to respond in kind as they selected their rating on the Stanford Sleepiness Scale. If the latter explanation were true, drivers would in effect be saying to themselves, "If I've been driving for a long time, then I must be tired."

Self-ratings did not correlate significantly with trip segments ranked according to percent of night driving, even though performance measures showed significantly reduced performance at night than during the day. If the "expectation" explanation of driving self-ratings in the previous paragraph is correct, a disturbing corollary would be that drivers had no expectation that night driving would be associated with reduced performance, when in fact these performance reductions are significant. Education concerning circadian influences on performance could be expected to provide these drivers with potential benefits.

INDIVIDUAL DIFFERENCES IN DRIVER SUSCEPTIBILITY TO DROWSINESS

There were large individual differences among drivers in levels of alertness and performance. These differences were generally consistent across different measures and were often apparent across comparable times-on-task, hours-of-day, and cumulative trips.

There was also a wide variation in the total number of episodes judged drowsy in the video records. Thirty-six percent (36%) of the drivers were never judged drowsy; of the remainder, 77% (49% of the total) were judged drowsy 10 or fewer times and 23% (15% of the total) were judged drowsy more than 10 times. Among the drivers with more than 10 drowsiness episodes, the number of drowsy episodes ranged from 12 to 40, with an average of 22 episodes during their 4-5 day participation period.

A further illustration of the wide individual differences among drivers is the fact that 11 of the 80 drivers (14%) accounted for 54% of all observed drowsiness episodes.

This study did not track the subjects over extended periods of time to determine if the same drivers showing frequent drowsiness during the week of the study would show frequent drowsiness weeks or months later. Thus, it cannot be discerned whether the observed individual differences were reflective of driver traits (i.e., long-term, stable individual differences in physiology and/or performance) or of driver states (short-term differences related to recent sleep or other transient events). Of course, both traits and states may be operative. Future research should address the trait versus state issue because it has implications for the potential effectiveness of improved driver selection, scheduling, and training as fatigue countermeasures.

SLEEP APNEA

PSG analysis of subject sleep revealed that 2 of the 80 drivers (2.5%) had clinically diagnosable sleep apnea. Interestingly, driving performance of these two individuals was not statistically different from that of other comparable drivers in the study, which may be because their total sleep time and sleep architecture variables were not grossly different from those of the other older drivers of Condition C1-10day.

AGE AND FATIGUE

No significant relationships were found between driver age and fatigue. There were no consistent differences between older and younger drivers in terms of observed drowsiness, frequency of naps, self-ratings, or driving performance. Older drivers performed more poorly on the Code Substitution test than younger drivers, but this effect was not fatigue-related. As shown by prior research, measures of perceptual-motor and memory abilities are generally lower for older subjects. In order to control for this effect, the Code Substitution data were disaggregated by driver age so that the general age difference in performance did not confound other comparisons.

FATIGUE COUNTERMEASURES

Numerous potential countermeasures to fatigue and loss of alertness have been reviewed in one way or another during the course of this study. In the literature review, countermeasures were divided into two major categories, having to do with (1) maintaining alertness at high levels, and (2) detecting and reacting to loss of alertness. It was concluded that the most important contributor in the first category is adequate sleep, particularly in the challenging but not uncommon case where the work-rest cycle is counter to the driver's natural body rhythms. Adequate sleep requires that drivers be provided adequate opportunities for sleep (including naps) *and* it requires that drivers place high priority on using available off-duty time for obtaining adequate sleep.

However, no scheduling efforts or hours-of-service rules can ensure that some drivers won't sometimes have loss of alertness to a degree incompatible with the goal of a crash-free environment, as set out at the 1995 Truck and Bus Safety Summit. Then it becomes critical for the impaired driver to recognize the danger and stop driving. Existing Federal Motor Carrier Safety Regulations require that the fatigue-impaired driver stop driving (49 CFR 392.3), but it doesn't always happen (Hayworth et al., 1988; NTSB, 1990 and 1995; McCart, 1995). The desired outcome in this acute situation requires (1) driver awareness of the impairment and (2) motivation to stop driving.

Timely detection of impairment may be improved by new driver education and training programs and by emerging technology-based off-line performance tests and real-time driver monitoring systems. Adequate driver motivation to stop driving when impaired requires reward, or at least minimum perceived cost, for doing so.

The optimum fatigue countermeasure system requires shippers, consignees, motor carriers, drivers, and regulators to promote effective driver training concerning fatigue, to accept proven technological innovations, and to implement fatigue-safe work schedules (including tolerance of occasional unscheduled nap stops) without excessive driver penalties in terms of dollars, assignment of subsequent work, or other consequences to a degree that suppresses the desired driver behavior.

It was beyond the scope of this study to specify in detail how this milieu might be achieved or to calculate all its costs, benefits, and other impacts. But the scientific data collected during this study clearly showed that among thousands of hours of notably professional and accident-free driving, there were some hours of extreme driver fatigue in the operational environment, inadequate driver self-detection of acute loss of alertness, and inadequate motivation to stop driving when experiencing it. Every such instance in the future will represent a risk -- and an opportunity to improve highway safety.

Assessment of Surrogate Performance Tests for Application to Fitness-For-Duty Testing

One purpose of this study was to assess the usefulness of automated measures predictive of driver fatigue. The periodic surrogate performance testing reported here represents one approach, namely fitness-for-duty testing prior to commencing a period of driving. The three performance tests employed were chosen because prior research indicated they measured abilities that are both fatigue and safety related. For example, the Simple Response Vigilance Test (SRVT) measures both vigilance and simple response latencies, and alertness and rapid response are clearly needed for adequate response to unexpected highway events.

The two SRVT measures (LAPSES and RVS) consistently showed statistically significant degradation associated with the length of time since the start of the trip, and with the number of trips completed. These performance losses are consistent with hypothesized acute and cumulative fatigue.

The sensitivity of the SRVT measures is made all the more remarkable by the dilution of variance from an uncontrolled nuisance variable associated with varying levels of sunlight. Because of their simplicity, face validity, and sensitivity, the SRVT and similar simple response tests may prove to be feasible predictors of fatigue in transportation settings. However, investigators must address, among other issues, the problem of uncontrolled variation in stimulus intensity associated with lower contrast ratios in sunlight, as well as the need to develop concrete pass/fail criteria.

The Code Substitution Test (CS) was found to be sensitive to acute performance losses at the end of trips and during the night. The large CS improvement from one trip to the next reflected ongoing skill acquisition. Some drivers were still improving CS performance after 20 distributed three-minute sessions. The test measures choice response abilities not required in the SRVT, making CS a good complement to SRVT that may well be more sensitive to acute fatigue. Additionally, CS did not show performance degradation in sunlight, a quality that recommends its use in situations where it would be difficult to control varying light levels.

Simple reaction time increases with decreasing stimulus intensity, whether the stimulus is auditory or visual. This is a likely explanation of the increased SRVT response times we found in daylight and in the ten-hour Conditions. The Code Substitution test, however, is not a vigilance and simple response task like the SRVT, but is a complex choice response task with components of visual searching, short-term visual memory, short-term motor memory, symbol recognition, and choice response speed. It is much closer, in terms of abilities exercised, to the reading-comprehension task employed by Garcia and Wierwille (1985) in their study of the effect of glare on a video display. Garcia and Wierwille failed to find a statistically significant relation between glare and correctness of responses on their reading task. In their study, glare increased the amount of time needed to read text judged easy to read, but decreased the amount of time spent on more difficult text.

The Critical Tracking Test (CTT) failed to show statistically significant changes associated with time of day, hours since trip start, or number of trips completed. Although this test has a long history of use in fatigue and driving research, the lack of correlation with the independent variables in this study raises questions about its utility as a measure of fatigue in future studies. The study, however, observed a small but statistically significant correlation of CTT with driver self-rating on the Stanford Sleepiness Scale ($r=-0.12$, $p< .001$).

Small, but statistically significant, correlations were also found between the driver's self-rating on the Stanford Sleepiness Scale and RVS, and LAPSES. The small size of the correlations suggests that drivers are only minimally aware of loss of the abilities measured by the performance tests, or it may be that the Stanford Sleepiness Scale is poorly suited for use by drivers in reporting safety-related performance losses they are aware of.

Implications of Study for Driver Education Concerning Fatigue Effects and Self-Awareness

Two major project findings relevant to driver education were the generally inadequate amounts of sleep obtained by the driver subjects and the strong tendency for drowsiness to be most associated with night driving. Drivers need to be educated to obtain enough sleep (about 7 hours according to self-reports by the drivers in the DFAS), especially if they need to drive at night. Further, study findings showed that drivers were generally poor judges of their own levels of fatigue/alertness based upon comparisons of their self-assessments with objective measures of their performance. This indicates a need to train drivers to better assess their current levels of fatigue while driving, perhaps by learning to become more conscious of changes in their physical state and subtle changes in their driving performance.

Summary of Study Findings Concerning Countermeasures

Although the DFAS was not designed specifically to support the development of technological countermeasures, the findings of the study are supportive of their feasibility.

Changes in driving performance, measured by increased variability in steering and lane tracking, were shown to be correlated with drowsiness as judged in video observation of the driver's face. The correlation between drowsiness and degraded driving performance supports the concept of continuous monitoring of driver performance to detect fatigue. A related and complementary approach to performance monitoring is to directly measure psychophysiological changes such as the eyelid droop seen in face videos of drowsy drivers or various PSG indices of reduced alertness. The DOT and other agencies and organizations are sponsoring a wide range of research on technological fatigue-detection and prevention countermeasures. Fitness-for-duty (readiness-to-perform) testing and continuous driver monitoring approaches are both being assessed. At the same time, driving performance is also influenced by the design and condition of the roadway, by the characteristics of the vehicle being driven, and by the number and location of other vehicles sharing the roadway. Those influences must be accounted for in the development of continuous-monitoring systems.

The study also sought to identify any behavioral methods used by drivers to ward off fatigue. Napping was a frequent driver-initiated response to drowsiness and fatigue. The alertness-enhancing effects of napping have been demonstrated in other operator performance settings (e.g., aviation) and should be the subject of future research, including re-analysis of the DFAS database. Other methods will be reported in a countermeasures survey to be published by the FHWA in the near future.

The powerful time-of-day effects on fatigue demonstrate that scheduling may be used as an important countermeasure to CMV driver fatigue. From the driver fatigue and alertness standpoint, the optimal schedule is one that appropriately manages night driving. There are no known highway transportation hours-of-service regulations in the world that address time-of-day effect, even though shiftwork literature for many years has pointed out a strong relationship between time-of-day and accidents and incidents.

It cannot be concluded, however, that shifting truck traffic to daylight hours would result in lower accident rates. This measure would increase daytime traffic congestion, possibly with a corresponding increase in accidents, and would further increase the risk of accidents with passenger vehicles which are more vulnerable in accidents with trucks because of their difference in mass. Research is needed to establish the relative risks of collision between day and night driving for a variety of road and vehicle types, and levels of traffic density, so as to establish the net impacts on highway safety of day/night scheduling practices.

Another key to improved scheduling at the fleet level may be the finding of large individual differences in susceptibility to drowsiness while driving. Some drivers may be much better than others at maintaining alertness in the long-haul CMV environment, especially at night -- a potential basis for driver selection and assignments of runs.

RECOMMENDATIONS

Table 8 summarizes recommendations in the context of the taxonomy of fatigue countermeasures developed during this study.

Table 8. Recommendations in context of fatigue countermeasures taxonomy

FATIGUE COUNTERMEASURES			
ALERTNESS MAINTAINERS		ALARMS	
Driver Controlled	Externally Controlled	Performance Based	Physiology Based
• Train drivers regarding sleep	• Study selection testing for night drivers	• Expand development of performance-based fitness-for-duty tests	• Expand research on relationship of driver sleep quality and circadian rhythm to performance
• Study ways to improve self-detection of fatigue	• Schedule in accordance with scientific knowledge of fatigue, performance, & sleep	• Expand development of real-time driver monitors/ warning systems	•Support research on drowsiness detection while driving
• Study effects of naps in commercial driving	• Hours of Service Rules		

DRIVER-CONTROLLED ALERTNESS MAINTAINERS

The first column of Table 8 deals with countermeasures drivers themselves can take to maintain alertness.

Train Drivers Concerning Sleep

Many of the drivers who participated in the data collection did not put as high a priority on obtaining sleep as they should have. Drivers (and most other nonspecialists) are not very aware of the considerable body of knowledge that has been accumulated recently regarding sleep and work scheduling. Consequently, there is good potential for improved drivers' sleep hygiene as a result of appropriate training. Suitable training programs and materials should be developed.

Improve Self -Detection of Fatigue

The sensitivity and reliability of driver self-detection of acute fatigue can probably be improved. It is likely that safety payoffs can be achieved earlier in this area than can be obtained with the longer lead time technology-based fatigue detectors. Self-detection and automated detection, of course, will be complementary. This issue should be investigated.

Use of Napping as a Countermeasure

Naps have been assessed as potential fatigue countermeasures in aviation and in other occupational contexts, as reported earlier. The effects of napping in commercial driving should be investigated, based in part on data collected during the naps taken in this study. The logical action for the driver who detects an acute loss of alertness in himself is to stop driving and nap. Drivers might also take prophylactic, or precautionary, naps prior to commencing a driving period. Naps during the mid-afternoon “siesta” period should be emphasized for drivers with duty start times between about 1500 and about midnight. Those behaviors likely can be facilitated and optimized with additional research leading to focused recommendations for commercial drivers.

EXTERNALLY CONTROLLED ALERTNESS MAINTAINERS

The second column of Table 8 deals with maintenance of alertness through action external to the driver.

Study Testing Methods for Selection of Night Drivers

Some of the 80 drivers who participated in the data collection experienced much more drowsiness than the others. Research should be conducted to determine if there are stable differences among drivers in this regard and to assess the relative contributions of individual trait and situational state. Practical screening tests might then be developed to identify persons not well suited to night driving.

Improve Scheduling Practices

Motor carriers should endeavor to schedule drivers in a manner which takes account of the now extensive scientific knowledge of the relationships among fatigue, performance, and sleep. Scheduling guidelines should be developed for use in the industry, and more motor carriers should begin to use this information for scheduling drivers.

Research Sleep Needs and Circadian Rhythm Issues as They Apply to Hours-of-Service Regulation and Education Programs

The hours-of-service rules are one means of externally controlling drivers to help prevent unsafe levels of fatigue. However, the hours-of-service rules don't reflect current knowledge of work-rest scheduling, especially concerning circadian rhythms.

Additional study should be conducted to establish a quantitative relationship between total sleep time, hours since awakening, and probability of drowsiness during vulnerable periods of the body's daily cycle. This would help to establish amounts of sleep and periods of performance that would provide safe levels of performance during vulnerable periods. This knowledge would be valuable to drivers, motor carriers, and policy makers. The program should ultimately use experienced professional drivers as subjects, since it is reasonable to assume that some drivers may self-select based on ability to tolerate irregular schedules.

Because of the risk associated with conducting such research on public highways, the use of simulators and test tracks should be evaluated with the specific commercial motor vehicle research goals in mind.

DEVELOP PERFORMANCE-BASED WARNING SYSTEMS

The third column of Table 8 concerns performance-based alarms or detection systems.

Fitness-for-Duty Tests

Fitness-for-duty tests can be used prior to going on duty, and on board the commercial vehicle as well, to aid the driver in assessing his status periodically.

From the available data, it appears that vigilance and reaction time tests such as the Simple Response Vigilance Test (SRVT) have utility in assessing a driver's condition on two critical safety-related abilities: vigilance and psychomotor response times. Both abilities are required for adequate response to unexpected highway events. Whether adequate and appropriate response is taken can often determine whether an accident occurs or is avoided.

Both the Code Substitution and Simple Response Vigilance Tests were able to quantify safety-related performance losses associated with fatigue. Though they show promise as fitness-for-duty tests, each also poses a challenge to practical implementation in the transportation workplace.

The Code Substitution Test taps cognitive abilities not measured by the Simple Response Vigilance Test. It has the advantage of relative immunity to variation in ambient light level, allowing use in environments where light level is not easily controlled. Additionally, meaningful scores can be obtained with a relatively short duration (180 seconds or less) test. However, the test has the disadvantage of relative complexity and long driver training period. Many of the drivers in the current study were still showing practice effects after 20 distributed three-minute test administrations. Code Substitution is recommended for future development leading to application in the workplace. Users must, however, take adequate steps to ensure that drivers are given adequate practice before use of test results in fitness-for-duty decisions.

The SRVT has the advantage of simplicity, making the test easy to learn. The simplicity of the SRVT also facilitates the engineering of a relatively economical testing device. However, SRVT scores are quite sensitive to variation of ambient light level, which makes the in-vehicle deployment of practical devices difficult. In addition, the test is relatively long (10 minutes), making test-taking onerous. The SRVT is recommended for further development with the goals of shortening test duration and decreasing sensitivity to ambient light level.

It must be stressed that the greatest variance in the performance test data was attributable to differing abilities across individuals. Any use of these tests for fitness-for-duty testing would require comparison of each driver's state to his norms. Furthermore, since the tests can only assess an individual's performance at the time they are taken, their role in an alertness- and fitness-management setting must be carefully assessed.

Real-Time Monitors and Warning Systems

Real-time loss-of-alertness monitors with appropriate characteristics (high probability of detection, low false alarm rate) are technically challenging, high-payoff projects which, when successful, will offer a very important enhancement to self-detection of fatigue. The data acquired in this study should be useful in moving both types of countermeasures toward their goals.

As has been noted in other studies, driving performance, as measured by lane tracking and steering behaviors, was affected by the level of alertness of the driver. However, because drivers alter their driving behaviors in response to conditions in the roadway environment, it is important to account for those differences. For example, the lateral position of a vehicle can change according to the traffic conditions in the adjoining lanes, as well as in response to the condition of the roadway itself (ruts in the wheel path, the cross-slope of the road, and other factors). Additionally, sensors that monitor the vehicle's position with respect to a fixed mark, such as a lane stripe, operate under the assumption that the mark is both properly placed and is readable by the sensor. These issues should be studied.

One of the most sensitive indicators used in this study was observation of drivers on video recordings. This allowed moment-by-moment evaluation of the driver's apparent state of alertness. Cues like eyelid behavior could prove very important for real-time monitors if reliable and affordable on-board automated measurement can be achieved. This is a very promising area for research.

SUPPORT PHYSIOLOGY-BASED METHODS

The fourth column of Table 8 addresses physiologically based detection of fatigue impairment.

Undertake Research Concerning Relationships Among Sleep Quality, Circadian Rhythm, and Performance

The present study benefited greatly from monitoring the principal sleep periods of the drivers, but there is more to be learned. For example, studies planned and in progress of sleep taken in sleeper berths and sleep taken on varying driving schedules in short- and long-haul operations where a wide range of total sleep times are obtained should lead to expanded knowledge with practical applications. These and other studies must take into account circadian rhythms in sleep quantity and quality and in human performance.

Undertake Research on Use of Physiological Measures for Drowsiness Detection

The collection and processing of physiological data for the purpose of supporting real-time drowsiness detection research are also important, although it does not appear from this study that current polysomnographic methods will be as sensitive to drowsiness or as practical to apply in operational settings as other indicators. In a research program, however, the physiological data can prove very valuable, particularly when assessed in conjunction with the behavioral data.

FUTURE DIRECTIONS

The DFAS demonstrated that it is possible to conduct a field study with substantial numbers of commercial drivers (80), hauling revenue freight on many trips (360), employing instrumentation to record numerous aspects of vehicle control, surrogate test performance, driver physiology, face and road video, and sleep studies for each principal sleep period. This was accomplished without any motor vehicle crashes or other harm to the drivers or other participants. The data collected have been used to document a number of fundamental characteristics of driver fatigue and alertness and the archived DFAS database will continue to support analyses of additional questions over the coming years.

Driver drowsiness/fatigue has become the dominant human factors research issue relating to CMV transportation. For example, the FHWA/OMC currently sponsors more than a dozen research and education/outreach projects relating to CMV driver drowsiness/fatigue. Recently completed, current, or planned fatigue-related research includes studies on work, rest and recovery; sleep apnea; multi-trailer vehicle driver stress and fatigue; highway rest areas; use of on-board recorders for hours-of-service and other regulatory compliance; driver fitness-for-duty testing; other technological fatigue countermeasures; fatigue education for driver and other CMV-related personnel; fleet-based driver wellness; the role of shippers and receivers in HOS violations; scheduling practices and fatigue; safety analysis of HOS "restart" options; sleeper berth usage and fatigue; local/short-haul driver fatigue; loading/unloading and fatigue; and improved crash causation analysis. Most of these studies will build upon the research techniques developed, practical lessons learned, and scientific knowledge gained from the DFAS.

A separate report, to be published under the sponsorship of Transport Canada and the Canadian Trucking Research Institute, will present the results and analysis of an additional field study that was performed in coordination with this one. The database in this study covers an additional 55 trips performed under various driving and days-off schedules lasting up to 10 days. That study will provide additional information to build on the results of this study, including results on the influence of different durations of multiday off-duty periods.