Human Factors: The Journal of the Human Factors and Ergonomics Society
http://hfs.sagepub.com/

Impairment of Driving Performance Caused by Sleep Deprivation or Alcohol: A Comparative Study

Stephen H. Fairclough and Robert Graham

Human Factors: The Journal of the Human Factors and Ergonomics Society 1999 41: 118
DOI: 10.1518/001872099779577336

Impairment of Driving Performance Caused by Sleep Deprivation or Alcohol: A Comparative Study

Stephen H. Fairclough and **Robert Graham**, HUSAT Research Institute, Loughborough University, Loughborough, Leicestershire, United Kingdom

A study was conducted to assess the relative impact of partial sleep deprivation (restriction to 4 h sleep before testing) and full sleep deprivation (no sleep on the night before testing) on 2 h of simulated driving, compared with an alcohol treatment (mean blood alcohol content = 0.07%). Data were collected from the 64 male participants on the primary driving task, psychophysiology (0.1 Hz heart rate variability), and subjective self-assessment. The results revealed that the full sleep deprivation and alcohol group exhibited a safety-critical decline in lane-keeping performance. The partial sleep deprivation group exhibited only noncritical alterations in primary task performance. Both sleep-deprived groups were characterized by subjective discomfort and an awareness of reduced performance capability. These subjective symptoms were not perceived by the alcohol group. The findings are discussed with reference to the development of systems for the online diagnosis of driver fatigue. Potential applications of this research include the formulation of performance criteria to be encompassed within a driver impairment monitoring system.

INTRODUCTION

Driver impairment is defined as a change in driving performance caused by the deleterious effects of alcohol, fatigue, drugs, or sudden illness (Fairclough, 1997a). In 1995, approximately 24% of traffic casualties in the United Kingdom were driving under the influence of an illegal level of alcohol – that is, > 0.08% blood alcohol content (BAC; Clayton, 1997). A recent survey of UK drivers indicated that driver fatigue was a contributory factor in 9% to 10% of accidents (Maycock, 1995).

The development of transport telematics systems has been proposed as a potential countermeasure to accidents induced by driver impairment. These systems encompass real-time monitoring, diagnosis, and feedback of driving impairment (e.g., Brookhuis, De Waard, & Bekiaris, 1997; Haworth & Vulcan, 1991; Mackie & Wylie, 1990; Wierwille, 1994). The goal of this technology is to provide predictive feedback concerning early symptoms of impaired driving. It is predicted that warning feedback may persuade the driver to break from the journey and, therefore, avoid those safety-critical episodes of impairment that increase the risk of accidents. Driver impairment monitoring technology relies on the sensitivity and validity of sensor apparatus in order to function effectively. The suitability of various sensor technologies may be assessed with reference to a number of measurement criteria (O'Donnell & Eggemeier, 1986): (a) that measures should be sufficiently sensitive to the earliest symptoms of impairment; (b) that measures are diagnostic – that is, capable of discriminating the influence of fatigue from other categories of impairment, such as alcoholic intoxication; and (c) that measures are selective and therefore able to distinguish the impairment "signal" against a dynamic and highly variable "noise" from the

Address correspondence to Stephen H. Fairclough, HUSAT Research Institute, Loughborough University, The Elms, Elms Grove, Loughborough, Leicestershire, LE11 1RG, United Kingdom; s.h.fairclough@lboro.ac.uk. **HUMAN FACTORS,** Vol. 41, No. 1, March 1999, pp. 118–128.

driving environment. It is argued that a combination of these criteria should be used to index the global accuracy of an impairment monitoring system.

Psychophysiological measures may be collected on a remote basis (i.e., there is no requirement to attach electrodes to participants); for instance, machine vision apparatus may be used to monitor eyelid activity (Tock & Craw, 1992). Although psychophysiology appears to be sufficiently sensitive, problems of selectivity are anticipated because of the inherently "noisy" environment inside the vehicle cockpit (e.g., fluctuations in temperature and lighting conditions). Behavioral measures are also potentially useful. However, it is postulated that certain symptoms, such as head-nodding when fatigued, occur at a relatively late stage of impairment and therefore may lack sufficient sensitivity. (Evidence for this view was provided by Haworth & Vulcan, 1991.)

The measurement of driving performance represents a valid strategy to index impairment (i.e., impairment is inferred directly on the basis of primary task performance rather than via a proxy measure). However, a number of problems are associated with the use of driving performance as a predictive and diagnostic source of impairment (Fairclough, 1997b). Specifically, primary task measures are deemed to be insufficiently sensitive to the presence of impairment; that is, task performance is protected from the influence of impairment by compensatory strategies (De Waard, 1996; Hockey, 1997). In addition, there is evidence of selectivity problems for performance measures. For example, in a study of simulated driving, Desmond and Matthews (1997) demonstrated that the presence of road curves was sufficient to suppress impaired vehicular control. Studies encompassing driver impairment caused by alcohol and fatigue have revealed a significant degree of overlap, with both showing effects of a reduction in the fidelity of responses to speed changes of a lead vehicle and an increase in the variability of lateral control (De Waard & Brookhuis, 1991). Therefore, it may be argued that driving measures are not sufficiently diagnostic to differentiate one category of impairment from the other.

Despite these problems, primary task measures occupy a pivotal role in the development of driver impairment monitoring systems. These measures form the main source of data for those telematic systems that employ multidimensional assessment (Brookhuis et al., 1997). In addition, the standard means of estimating the validity of indirect measures such as psychophysiology is to utilize changes in primary driving performance as a reference variable (Wierwille, 1994).

The aim of the current study was to assess changes in driver performance caused by the influence of two sources of impairment (sleep deprivation and alcoholic intoxication) within a simulated environment. The study included multidimensional measurement in order to assess the sensitivity of the primary task to changes in subjective self-assessment and psychophysiology. The inclusion of two different impairment manipulations within the study was intended to reveal the diagnosticity of primary measures. In addition, the simulated environment contained a number of driving scenarios (i.e., interactions with other vehicles) in order to allow us to index the selectivity of driving performance measures.

METHOD

Design

The study involved four between-subjects treatment groups, as follows: a control group (participants had a full night of sleep before the trial and did not receive alcohol), a partial sleep deprivation group (PartSD; participants were instructed to sleep for 4 h between midnight and 4:00 a.m. on the night before the trial and did not receive alcohol), a full sleep deprivation group (FullSD; participants were instructed to remain awake throughout the night before the trial and did not receive alcohol), and an alcohol group (participants had a full night of sleep on the night before the trial and received an alcoholic drink – a mixture of vodka and lemonade – before the experimental session). The amount of alcohol administered was calculated according to the Widmark equation as described by Walls and Brownlie (1985). Participants received an amount to approximate a peak level of BAC in the range of 0.08% to 0.1%. The trial consisted of three cumulative 40-min periods of simulated driving. The manipulation of driving scenario formed a second within-subjects factor and was designed to induce differing degrees of task demand.

This experimental design contained a degree of deception – that is, participants in the alcohol group were not briefed regarding the presence and volume of alcohol received. Therefore, the full procedure for the proposed experimental design was submitted, reviewed, and approved on ethical grounds by the Ethical Advisory Committee at Loughborough University. This committee raised a number of concerns. They recommended gathering more detailed information on participant health as well as highlighting and reminding participants of their right to withdraw from the study at any point without incurring a financial penalty.

Participants

There were 64 men participating in the experiment. All participants had normal or corrected 20/20 vision. Each participant was allocated to one of four experimental groups, each containing 16 participants. The groups were balanced according to the demographic variables of age, driving experience, average alcohol intake, and average hours of sleep per night across the four groups. Descriptive statistics for all four groups are presented in Table 1. None of the demographic variables was statistically different among the four groups. All participants were paid.

Apparatus

The experiment was performed using a fixed-base driving simulator. Participants were seated inside a Ford Scorpio and viewed a large projector screen (approximately 3 m × 4 m). A Pentium PC was used to simulate the vehicle model and to generate the driving scene. The computer-generated scene was projected on the screen by a Sony Multiscan projector. Unfortunately, the simulator software was not capable of providing interactive sound, but participants were exposed to a recording of in-vehicle sound collected from a real vehicle. Electrocardiographic (ECG) data were collected by an analog-to-digital converter and bioamplifiers connected to a Macintosh Powerbook™ computer.

Experimental Task

The simulated driving scene consisted of a straight, two lane, left-side-drive road under daytime conditions. The simulated road was flanked by vegetation and marker posts. Each 40-min block contained six driving scenarios: (a) open road (no other traffic), (b) following (a lead vehicle in the left lane traveling at a steady speed of 96.5 km/h, or 60 miles/h),

TABLE 1: Demographic Characteristics of the Four Groups (*N* = 64)

Participant Group	Age (years)	Driving Experience[a]	Average Duration of Sleep[b]	Average Alcohol Intake[c]
Control	30.63 (20–46)	12.8 (2–25)	7.7 (6.0–8.0)	16 (3–23)
PartSD	30.63 (22–47)	9.94 (2–2)	8.0 (7.0–9.5)	17 (4–28)
FullSD	30.63 (20–50)	11.3 (4–23)	7.7 (6.5–9.0)	14 (3–24)
Alcohol	30.68 (20–50)	12.5 (3–25)	7.8 (7.5–9.0)	16 (3–30)

Note: Minima and maxima given in parentheses.
[a] Number of years participants held a full driving license. [b] Self-assessed to the nearest 30 min. [c] Self-assessed units of alcohol consumed per week.

(c) passing (vehicles overtaking in the right lane at a rate of one every 50 s, (d) following/passing (a combination of Scenarios b and c), (e) low sinusoidal following (a lead vehicle in the left lane varying its speed sinusoidally between 88.5 and 104.6 km/h, or 55 and 65 miles/h, with a cycle time of 30 s; see Brookhuis, De Waard, & Mulder, 1994), and (f) high sinusoidal following (the lead vehicle varying its speed sinusoidally between 80.5 and 112.6 km/h, or 50 and 70 miles/h, with a 30-s cycle time). The presentation order of the scenarios was fixed within each 40-min block but was different among the three blocks. Each 40-min block began with a 10-min open-road scenario in which no data were collected.

Experimental Measures

Measurement of the driving task was divided into lateral control, longitudinal control, and speed control. Lateral control was measured at three levels of criticality: (a) frequency of "accidents," defined as those occasions when the vehicle left the road edge or when the participant collided with another vehicle; (b) frequency of lane crossings (when two vehicle wheels made contact with the left lane edge or the right lane boundary); and (c) frequency of near-crossing incidents when the minimum time to line crossing (TLC; Van Winsum, Brookhuis, & De Waard, 1998) fell below 2 s. In order to index participants' steering input, the frequency of steering wheel reversals was measured (the number of zero-crossings of the steering wheel, as defined by McLean & Hoffman, 1975). Longitudinal control was indexed via mean time headway (intervehicle distance divided by speed), and speed variability was calculated. All measures of primary driving performance were originally sampled at 10 Hz.

ECG data were collected via three disposable electrodes attached to the participant's chest. R-peaks of the ECG trace were detected and corrected for artifacts. The interbeat interval (IBI) variable was subjected to spectral analysis and decomposed into three bandwidths (Mulder, 1979) to extract the midfrequency band (0.07–0.14 Hz). This frequency band is related to short-term blood pressure regulation and has been successfully employed to index driver mental workload (De Waard, 1996).

The study employed a battery of subjective measurement questionnaires. Participants completed multiple administrations of (a) the University of Wales Mood Adjective Checklist (Matthews, Jones, & Chamberlain, 1990), (b) the NASA-Task Load Index calculated on the basis of raw ratings (RTLX; Byers, Bittner, & Hill, 1989; Hart & Staveland, 1988), (c) the Karolinska Sleepiness scale (Kecklund & Akerstedt, 1993), (d) a cognitive interference scale (Sarason, Sarason, Keefe, Hayes, & Shearin, 1986), and (e) an eight-point sobriety scale devised for the study.

Experimental Procedure

The participants came to our institute on two occasions: a practice session, then a test session on the following day. The practice session involved a supervised training journey in the simulator that lasted for 20 min, followed by a 20-min baseline journey, during which psychophysiology and driving data were collected.

All participants were provided with transportation to and from the institute on the day of their test session. This arrangement was made to ensure the integrity of the alcohol placebo. On arrival, participants were provided with a drink. The control group and the two sleep-deprived groups received a placebo containing lemonade with a tablespoon of vodka floated on the surface to provide the smell of alcohol. Once the drink had been consumed, participants completed a 10-min familiarization journey. Following familiarization, all participants completed a pretest set of subjective questionnaires, and their breath was analyzed using an alcohol meter (Lion Laboratories Alcometer S-D2, Barry, Wales, UK). Participants then drove the first 40-min block of the simulated journey. On completion, participants filled out a second set of subjective questionnaires and their breath was analyzed once more. Both the subjective questionnaires and Alcometer tests were administered following completion of the second and third 40-min blocks. The recess after each journey block was restricted to 5 min. On completion of the test session, participants were debriefed, paid, and transported back to their homes.

Figure 1. Box and whiskers chart illustrating the percentage blood alcohol content levels recorded for the alcohol group across the test session ($N = 16$).

RESULTS

Alcohol Manipulation

Administration of alcohol using the Widmark equation is an inexact procedure; the means, standard errors, and standard deviations of actual BAC figures obtained during the study are shown in Figure 1. These data illustrate that the participants in the alcohol group were on a descending BAC curve through the experimental session, dropping from a mean of 0.08% to 0.05% across the 2-h journey.

Primary Task Performance

The raw vehicle data were averaged across a 5-min time window for each driving scenario, and data were analyzed using multivariate analysis of variance (MANOVA) techniques. When appropriate data were available from the baseline session and when the test contained a maximum of three main effects, multivariate analyses of covariance (MANCOVAs) were employed (i.e., using baseline data as covariates). If the test contained more than three factors, this caused a substantial problem of interpretation; therefore the covariate was not included. Within-subject significance was reported using the Wilks's lambda statistic, λ_w, to overcome problems of sphericity (Vasey & Thayer, 1987). All data were tested for the existence of outliers, defined as raw values further than 2.5 standard deviations away from the mean. All post hoc testing was performed via the Tukey HSD test.

In the course of the study, 51 accidents occurred. However, the distribution of accident frequency rendered data unsuitable for analysis – that is, 10 participants accounted for all of the accidents, and of these, 2 were responsible for 24 accidents. The distribution of these participants across the four experimental groups was as follows: 5 from the FullSD group, 1 from the control group, and 4 from the alcohol group.

The analysis of lateral control is presented with respect to two categories of data: those concerned with vehicle position and those capturing steering control. Descriptive statistics revealed the presence of an outlier in the lateral control data; therefore, 1 participant from the control group was not included in these analyses. All variables were subjected to a 4 (condition) × 3 (time on task) × 6 (scenario) MANOVA. A significant main effect for condition, $F(3, 59) = 4.58$, $p < .01$, revealed that the frequency of lane crossings was higher for the FullSD group ($M = 5.8$ per 5 min) and the alcohol group ($M = 4.8$), compared with the PartSD group ($M = 3.4$) and the control group ($M = 3.6$). There was a significant effect of time on task, $\lambda_w(2, 118) = .76$, $p < .01$; that is, the mean frequency of lane crossings rose by one per 5 min from the beginning to the end of the journey. The main effect of scenario revealed that the highest frequency of lane crossings occurred during the open-road and passing scenarios, $\lambda_w(5, 295) = .44$, $p < .01$. A significant interaction of Condition × Time on Task showed that the frequency of lane crossings was significantly lower in the alcohol group than in the FullSD group during the final 40 min of the journey, $\lambda_w(6, 118) = .79$, $p < .01$.

The frequency of near-lane-crossings was indexed by those occasions when TLC fell below 2 s – that is, when the lateral velocity of the vehicle meant that it was 2 s away from a lane crossing. These data revealed a marginal effect of condition, $F(3, 59) = 2.60$, $p < .06$, which indicated a higher frequency of near-lane-crossings for the PartSD group ($M = 4.6$ per 5 min) than for the other three groups.

Steering control was analyzed as the steering wheel reversal rate per minute. Multivariate analysis revealed a significant main effect for condition, $F(3, 59) = 26.6$, $p < .05$, indicating that mean steering wheel reversal rate was higher for the control and alcohol groups ($M = 15.6$ and 14.3, respectively) compared with the PartSD and FullSD groups ($M = 11.2$ and 10.9, respectively). A significant main effect for scenario, $\lambda_w(5, 295) = .56$, $p < .01$, revealed that levels of steering activity were

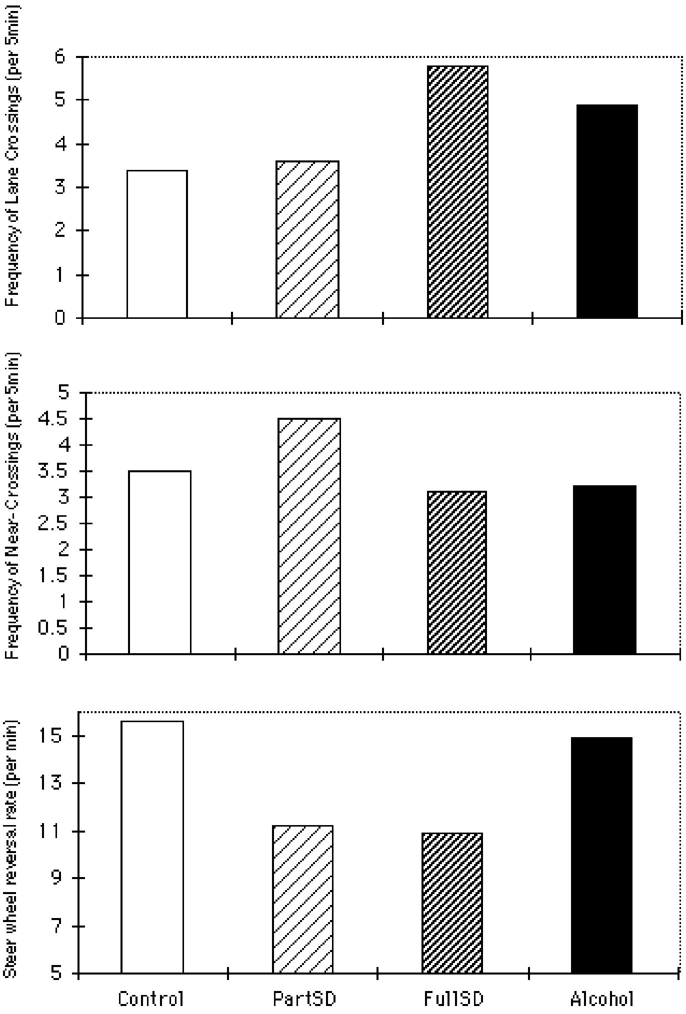

Figure 2. Lateral vehicle control as represented by mean frequency of lane crossings, mean frequency of near-lane-crossings, and steering wheel reversal rate across the four treatment groups ($N = 63$).

higher during the high-sinusoidal-following and following scenarios (M = 15.0 in both cases) than during the open-road scenario (M = 12.6). The main effects for all three measures of lateral control across the four experimental groups are represented in Figure 2.

Analysis of mean time headway was confined to those four scenarios in which a lead vehicle was present. Of the participants, 6 were omitted from this analysis because of accidents or because they merely lost sight of the lead vehicle. This analysis showed a marginal effect for condition, $F(3, 54) = 2.22$, $p = .08$. Post hoc tests revealed a significant difference between the alcohol (M = 3.10 s) and FullSD groups (M = 3.98 s). There was also a significant effect for time on task, $\lambda_w(2, 108) = .88$, $p < .05$; mean headway decreased by 0.5 s between the initial 40 min and the final 40 min of the journey.

Speed variability was measured as standard deviation of speed (across a 5-min time window). The analysis showed a significant main effect for condition, $F(3, 59) = 4.56$, $p < .01$. Post hoc testing revealed that speed variability was lower for the alcohol group (M = 6.7 km/h) than either the PartSD or FullSD groups (M = 9.2 km/h for both). In addition, there was a significant main effect for time on task, $\lambda_w(2, 120) = .89$, $p < .05$, providing evidence of an increase in speed variability over time.

Psychophysiology

It is known that alcohol has a confounding effect on mean IBI and the 0.1 Hz component of heart rate variability (Gonzalez-Gonzalaz, Llorens, Novoa, & Valeriano, 1992). Therefore, ECG data from the participants in the alcohol group were not subjected to statistical testing. In addition, ECG data were dropped from 7 participants across the three remaining groups because of measurement artifacts.

Raw IBI data were subjected to analysis using Carspan software (Mulder & Schweizer, 1993) to isolate the midfrequency component of heart rate variability. These data were subjected to a natural log transform and a baseline conversion prior to parametric MANOVA analysis. There were significant main effects for condition, $F(2, 32) = 2.52$, $p = .05$, and time on task, $\lambda_w(2, 64) = .64$, $p < .01$. It was apparent that the midfrequency component was significantly suppressed for the PartSD group compared with the control group ($p < .05$). In addition, mean power in the midrange frequency increased between the first and third periods of the journey ($p < .05$). These findings are indicative of a higher level of mental-effort investment for the PartSD group compared with the control participants.

Subjective Data

Participants completed the questionnaires on four occasions through the test session, including a pretest administration. These subjective data were analyzed in a series of 4 (condition) × 4 (time on task) MANCOVAs with data collected from the end of the practice session used as covariates. The results of these analyses are shown in Table 2. It should be noted that there were no significant interaction effects.

Measurement of subjective mood revealed a significant decline of energetic arousal for the two sleep-deprived groups compared with the control participants. In addition, sleep deprivation raised subjective workload via increased effort and reduced estimates of performance efficacy. The FullSD group experienced increased temporal demand and physical demand in comparison to the other three groups. The effect of alcohol was to significantly reduce the level of frustration experienced by participants as compared with the other three groups. It was found that the frequency of task-relevant thoughts was higher for the FullSD group than for either the control or alcohol group. The PartSD group exhibited a similar pattern, but frequency was significantly higher than only the control group. Subjective ratings of both sleepiness and sobriety were included to reference the experimental manipulations. Subjective sleepiness was significantly higher for both sleep-deprived groups compared with the control group (but did not differentiate between PartSD and FullSD). Self-rated drunkenness was highest for the alcohol group.

The time on task effects for subjective measures were all in the expected direction – that is, workload factors, task-irrelevant thoughts, and subjective sleepiness all increased across

TABLE 2: Significance Tests of MANCOVA Analyses of Subjective Data ($N = 64$)

	Condition $F(3, 60)$	p	Time on Task $\lambda_w(2, 120)$	p
Mood				
Energetic arousal	5.88	< .01	.54	< .01
Tense arousal	1.43	ns	.94	ns
Hedonic tone	1.34	ns	.65	< .01
Subjective mental workload				
Mental demand	1.09	ns	.92	ns
Physical demand	4.41	< .01	.50	< .01
Time demand	4.06	< .05	.95	ns
Performance	2.73	.05	.80	< .01
Effort	3.16	< .05	.87	< .05
Frustration	4.51	< .01	.61	< .01
Mean workload	2.63	.05	.68	< .01
Cognitive interference				
Task Relevant	4.23	< .01	.60	< .01
Task Irrelevant	1.29	ns	.45	< .01
Sleepiness				
Karolinska scale	11.37	< .01	.39	< .01
Sobriety				
Self-rating	12.11	< .01	.81	< .05

the trial duration, whereas energetic arousal, hedonic tone, and task-relevant thoughts all decreased.

DISCUSSION

The effects of sleep deprivation and alcohol on driving performance may be categorized in terms of safety-critical changes (i.e., driver errors likely to increase the probability of accident) and non-safety-critical changes (i.e., alterations in vehicular control that do not increase the probability of accident).

It was apparent that a full night without sleep was sufficient to impair those participants with respect to both categories. The FullSD group exhibited the highest frequency of lane crossings and the highest mean time headway separation to a lead vehicle. The former effect may have been the direct result of a reduced level of steering input, whereas the latter may have been an adaptive strategy to counteract an increased risk of rear-end collision.

The effects of a reduction of sleep were more paradoxical in the case of the PartSD group. These participants exhibited normal lateral control (relative to the control group) but functioned on the same reduced level of steering input that characterized the FullSD group. The only evidence of impairment for the PartSD group was an increased frequency of near-lane-crossings. These data suggested that PartSD participants exhibited a proactive strategy of lateral control that operated on reduced levels of steering input – that is, they made a steering correction only when absolutely necessary to compensate for the threat of an imminent lane crossing.

The measures that exhibited the highest sensitivity to sleep deprivation were non-safety-critical changes in performance, subjective self-assessment, and psychophysiology. It was notable that noncritical changes in performance, such as reduced steering input and increased speed variability, characterized PartSD participants as well as the FullSD group. If both groups are considered representative of a continuum of rising impairment, sleep deprivation makes an impact on non-safety-critical behaviors at a moderate level (represented by the PartSD manipulation) and extends its influence to safety-critical performance at higher levels (represented by the FullSD manipulation). That is, the compensatory response of the driver to rising impairment may protect safety-critical task components at the expense of those that are less safety critical (Hockey, 1997).

It is postulated that a compensatory response to sleep deprivation was triggered by increased subjective discomfort and an awareness of reduced performance efficacy, which accompanied operational fatigue (Table 2). These participants responded by mobilizing mental effort to counteract the influence of sleep deprivation. This response was demonstrated by the analyses of subjective workload ratings (i.e., increased effort and increased frequency of task-relevant thoughts) and the suppression of the midfrequency component of heart rate variability that characterized the PartSD participants. These findings suggest an effective compensatory strategy on behalf of the PartSD group, with an adequate level of safety-critical performance sustained throughout the trial. However, the deterioration of lateral control, combined with heightened mental workload and a moderate suppression of heart rate variability, indicates ineffective compensation from the FullSD participants.

A differential awareness of subjective discomfort was one feature that distinguished the sleep-deprived participants from the alcohol group. Aside from decreased sobriety, the alcohol group exhibited reduced frustration compared with the other treatment groups. There was no evidence for any subjective awareness of impairment. In contrast to their sleep-deprived counterparts, the impairment of performance exhibited by the alcohol participants was limited to safety-critical changes such as more-frequent lane crossings (relative to control and PartSD participants) and a reduction of mean time headway (relative to the FullSD group). A second feature that distinguished the FullSD and alcohol participants was an interaction with time on task. The results revealed that impairment caused by sleep deprivation tended to increase with time on task; that is, lane crossing frequency peaked for the FullSD group during the final 40 min. By contrast, the impaired lateral control caused by alcohol was stable across the journey, despite a descending alcohol absorption curve (Figure 1).

The six driving scenarios were included to induce a variable level of mental workload throughout the simulated journey. The presence of a lead vehicle was the crucial feature that differentiated the level of lateral control across all groups. The absence of a lead vehicle during open-road and passing scenarios caused an increased frequency of lane crossings. The improvement of lateral control associated with a lead vehicle may have been attributable to (a) a perceptual effect (i.e., the presence of the lead vehicle on the simulated view functioning as a perceptual cue), (b) a workload effect (i.e., the lead vehicle functioning as a potential hazard and demanding a higher level of attention and control), or (c) a combination of both effects. The analysis of psychophysiology provided no evidence of increased mental effort when a lead vehicle was present. However, steering wheel reversal rate was elevated for scenarios that included a lead vehicle (e.g., following and high sinusoidal following) compared with the open-road scenario. Therefore, the presence of the lead vehicle appeared sufficient to stimulate steering activity, if not to raise mental effort.

CONCLUSIONS

The analyses of experimental data carry a number of implications for those measurement criteria necessary for the development of driver impairment monitoring systems. The results indicated that range of driving variables should be included in a monitoring system and that variables should be identified either as early or late indicators of impairment (i.e., non-safety-critical vs. safety-critical indicators). In addition, the level of association or dissociation between multiple measures may be indicative of the compensatory response of the driver to impairment, and this pattern may be used to differentiate between fatigued and intoxicated drivers. These results suggested that a reliance on safety-critical indices of performance has fostered the misleading view that measures of primary performance are relatively insensitive and undiagnostic with respect to impairment.

The study did not reveal any problems of selectivity for primary task measures – that is, none of the driving scenarios exerted sufficient influence to mask the influence of impairment on performance. However, this conclusion must be treated with caution because the manipulation of driving scenario was limited to the presence of other vehicles on the simulated roadway. It is anticipated that more powerful environmental manipulations (e.g., road geometry, traffic density, and reduced visibility) may reveal selectivity problems of the type described by Desmond and Matthews (1997).

The methodology employed during the study raises a number of points that can inform future research. First, the declining alcohol absorption curve may not have represented a valid comparison with rising fatigue caused by time on task. An alternative approach may be to contrast the influence of ascending alcohol absorption with sleep deprivation or to introduce additional administrations of alcohol throughout the session to induce a stable level of BAC. Although the inclusion of a placebo condition is the standard means of testing alcohol and other drugs, this protocol introduced a possible source of confounding in this study. Whereas the sleep-deprived participants knew exactly how much sleep they had lost and, therefore, may have formulated particular expectations regarding their anticipated level of impairment, the alcohol participants performed “blind.” All participants in the alcohol group were aware of the presence of alcohol but were unable to gauge the quantity they had consumed and, therefore, were unable to devise expectations regarding anticipated impairment.

Finally, from an ethical standpoint, the experimental design demanded that testing occur within a simulated environment rather than on the real road. Although we hope that the qualitative contrast between impairment manipulations has ecological validity, it is difficult to generalize this comparison quantitatively from a fixed-base simulator to the real road. In practical terms, the methodology satisfied the requirement to assess relative change attributable to impairment within a specific testing environment. Such specificity is double-edged, however, and generalization of data outside this particular environment may be limited. It is suggested that one step toward answering this question would be a replication employing a closed-circuit test track.

ACKNOWLEDGMENTS

This work was funded by the Commission of the European Communities as part of research conducted on the System for effective Assessment of the driver state and Vehicle control in Emergency situations (SAVE) project (TR 1047) under the Transport Telematics program. The authors wish to acknowledge the assistance received from Jayshree Lakha, Peter Marsh, and Dick De Waard.

REFERENCES

Brookhuis, K. A., De Waard, D., & Bekiaris, E. (1997). Development of a system for the detection of driver impairment. In C. Mercier-Guyon (Ed.), *Proceedings of the 14th International Conference on Alcohol, Drugs and Traffic Safety* (Vol. 2, pp. 581–586). Annecy, France: Centre d’Etudes et de Recherches En Medecine Du Traffic.

Brookhuis, K. A., De Waard, D., & Mulder, B. (1994). Measuring driving performance by car-following in traffic. *Ergonomics, 37*, 427–434.

Byers, J. C., Bittner, A. C., & Hill, S. G. (1989). Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary? In A. Mital (Ed.), *Advances in industrial ergonomics and safety* (Vol. 1, pp. 481–488). London: Taylor & Francis.

Clayton, A. (1997). The end of the decline in drink driving in Britain? In C. Mercier-Guyon (Ed.), *Proceedings of the 14th International Conference on Alcohol, Drugs and Traffic Safety* (Vol. 3, pp. 1227–1231). Annecy, France: Centre d’Etudes et de Recherches En Medecine Du Traffic.

Desmond, P. A., & Matthews, G. (1997). Implications of task-induced fatigue effects for in-vehicle countermeasures to driver fatigue. *Accident Analysis and Prevention, 29,* 515–524.

De Waard, D. (1996). *The measurement of driver mental workload.* Groningen, Netherlands: Rijksuniversiteit Groningen.

De Waard, D., & Brookhuis, K. A. (1991). Assessing driver status: A demonstration experiment on the road. *Accident Analysis and Prevention, 23,* 297–301.

Fairclough, S. H. (1997a). *Methods and criteria for the detection of impaired driving* (Report D5.1/SAVE TR1047). Athens, Greece: Transport Research and Development International.

Fairclough, S. H. (1997b). Monitoring driver fatigue via driver performance. In Y. I. Noy (Ed.), *Ergonomics and safety of intelligent driver interfaces* (pp. 363–380). Mahwah, NJ: Erlbaum.

Gonzalez-Gonzalaz, J., Llorens, A. M., Novoa, A. M., & Valeriano, J. J. C. (1992). Effect of acute alcohol ingestion on short-term heart rate fluctuations. *Journal of Studies on Alcohol, 53*(1), 86–90.

Hart, S. G., & Staveland, L. E. (1988). Development of the NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam: North-Holland.

Haworth, N. L., & Vulcan, P. (1991). *Testing of commercially available fatigue monitors* (Report 15). Victoria, Australia: Monash University, Accident Research Centre.

Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology, 45,* 73–93.

Kecklund, G., & Akerstedt, T. (1993). Sleepiness in long distance truck driving: An ambulatory EEG study of night driving. *Ergonomics, 36,* 1007–1018.

Mackie, R. R., & Wylie, C. D. (1990). Countermeasures to loss of alertness in truck drivers: Theoretical and practical consequences. In M. Vallet (Ed.), *Le maintien de la vigilance dans les transport* [Maintaining vigilance during transport] (pp. 113–141). Bron, France: Paradigme.

Matthews, G., Jones, D. M., & Chamberlain, A. G. (1990). Refining the measurement of mood: The UWIST Mood Adjective checklist. *British Journal of Psychology, 81,* 17–42.

Maycock, G. (1995). *Driver sleepiness as a factor in car and HGV accidents* (Report 169). Crowthorne, UK: Transport Research Laboratory.

McLean, J. R., & Hoffman, E. R. (1975). Steering reversals as a measure of driving performance and steering task difficulty. *Human Factors, 17,* 248–256.

Mulder, G. (1979). Mental load, mental effort and attention. In N. Moray (Ed.), *Mental workload: Its theory and measurement* (pp. 327–343). New York: Plenum.

Mulder, L. J. M., & Schweizer, D. A. (1993). *Carspan: Cardiovascular experiments analysis environment.* Groningen, Netherlands: iec ProGamma.

O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology. In K. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of human perception and performance* (Vol. 2, pp. 42.41–42.49). New York: Wiley.

Sarason, I. G., Sarason, B. R., Keefe, D. E., Hayes, B. E., & Shearin, E. N. (1986). Cognitive interference: Situational determinants and traitlike characteristics. *Journal of Personality and Social Psychology, 57,* 691–706.

Tock, D., & Craw, I. (1992, September). *Blink rate monitoring for a driver awareness monitoring system.* Paper presented at the British Machine Vision Conference, University of Leeds, United Kingdom.

Van Winsum, W., Brookhuis, K. A., & De Waard, D. (1998). *Approximations of time-to-line crossing (TLC) for real time measurement during car driving* (Report TM-98-B008). Soesterburg, Netherlands: TNO Human Factors Institute.

Vasey, M. W., & Thayer, J. F. (1987). The continuing problem of false positives in repeated measures ANOVA in psychophysiology: A multivariate solution. *Psychophysiology, 24,* 479–486.

Walls, H. J., & Brownlie, A. R. (1985). *Drink, drugs & driving.* London: Sweet & Maxwell.

Wierwille, W. W. (1994, May). *Overview of research on driver drowsiness definition and driver drowsiness detection.* Paper presented at the 14th International Conference on Enhanced Safety of Vehicles, Munich, Germany.

Stephen H. Fairclough received a B.Sc.(Hons.) in psychology from Lancashire Polytechnic in the United Kingdom in 1986. He has worked at the HUSAT Research Institute since 1987 and is a senior research fellow at Loughborough University in the United Kingdom.

Robert Graham received a M.Sc. in work design & ergonomics from the University of Birmingham, England in 1992. He has worked as a human factors researcher and consultant at the HUSAT Research Institute since 1992.

Date received: February 18, 1998
Date accepted: August 4, 1998