

An On-Road Assessment of the Impact of Cognitive Workload on Physiological Arousal in Young Adult Drivers

Bryan Reimer*, Bruce Mehler, Joseph F. Coughlin, Kathryn M. Godfrey and Chuanzhong Tan
MIT AgeLab & The New England University Transportation Center
77 Massachusetts Avenue, E40-291
Cambridge, MA 02139

*Corresponding author: reimer@mit.edu (617) 452-2177

ABSTRACT

In this paper, we describe changes in heart rate and skin conductance that result from an artificial manipulation of driver cognitive workload during an on-road driving study. Cognitive workload was increased systematically through three levels of an auditory delayed digit recall (n-back) task. Results show that changes in heart rate and skin conductance with increasing levels of workload are similar to those observed in an earlier simulation study. Heart rate increased in a step-wise fashion through the first two increases in load and then showed a less marked increase at the highest task level. Skin conductance increased most dramatically during the first level of the cognitive task and then appeared to more rapidly approach a ceiling (leveling) than heart rate. Findings further demonstrate the applicability of physiological indices for detecting changes in driver workload.

Categories and Subject Descriptors

J4 Social and Behavioral Sciences: Psychology; J.7 Computers in Other Systems: Real time; H5.m Information Interfaces and Presentation: Miscellaneous

General Terms

Measurement, Performance, Experimentation, Human Factors.

Keywords

Detecting driver state, cognitive workload, driving performance, physiology, driver distraction.

1. INTRODUCTION

The increase in the complexity of information available to drivers through in-vehicle interfaces and handheld devices has the potential to increase driver workload to a level at which driving performance begins to decline, thus increasing accident risk [3, 16]. As the percentage of older drivers increases, the advent of autonomous driving systems, and continued surge of secondary demands, active workload managements systems will increasingly need to monitor and manage driver state [1, 12]. Examples of early generation systems include Volvo's Intelligent Driver Information System [5] and Saab's ComSense [13]. These systems primarily help drivers prioritize information and delay the presentation of dynamic content based on the driving situation. Future workload management systems may better adapt to changing demands by incorporating real time measures of individual drivers' capacity. Physiological measures are a noninvasive method of characterizing relative cognitive workload

[3] and have been suggested as being complementary to driving performance based measures in a more comprehensive assessment of driver workload [16]. Further, physiological measures have been frequently described as being potentially more sensitive to initial changes in workload than performance based measures [2, 7, 8, 15] although only limited published data is currently available to support this assertion [8]. In this research, we further explore the capability of physiological indices to discriminate subtle changes in driver workload.

2. BACKGROUND

Theoretically, performance and arousal have an inverted U relationship (Yerkes-Dodson law); performance increases with arousal up to an optimal point and then declines as workload and the arousal associated with it continues to build. This point or, more broadly, optimal operating range, will vary with differences in operator capacity and other individual and situational factors. Drivers too often function in under aroused or distracted states that result in suboptimal arousal levels and a higher potential for accidents [16]. Real-time detection of an operator's workload might be used to provide feedback that attempts to help the driver self-manage the demands of operating the vehicle and associated tasks, thereby optimizing workload and enhancing automotive safety [12].

In a series of driving studies, heart rate and other physiological indices have been shown to be responsive to increased cognitive demand [2, 4, 7]. However, the literature on the relationship between heart rate, other physiological indices, and driving performance appears to be mixed. Lennenman, Shelley & Black [7] found that simulated driving performance did not degrade with increasing workload and heart rate. Engström, Johansson & Östlund [6] reported changes in physiological measures that were inconsistent across dual task conditions, visual vs. cognitive, and between locations, simulation vs. field. In a simulation study in our lab, Mehler et al. [8], heart rate and skin conductance detected changes in workload prior to a discernable decrement in driving performance. A general decrease in driving performance was only apparent at the highest level of secondary task demand. In a field study using the same secondary task presentation, Reimer [11] showed that changes in visual attention appear to be another method of detecting levels of secondary cognitive workload. Consistent with the earlier simulation work [8], a decrease in driving performance was only found at the highest level of workload. In the present paper, we report on physiological data collected during this study.

3. METHODS

3.1 Participants

Twenty-six drivers between 22 and 27 years of age participated in the study ($M = 23.85$, $SD = 1.57$). Participants were required to read and sign an approved informed consent form, to present a valid driver's license, and attest to having their license for more than three years, driving more than three times per week and being in good health. A research assistant ensured that participants clearly understand and spoke English. Participants were excluded if they had been involved in a police reported accident in the past year, needed glasses to drive, or were taking a medication that caused drowsiness. Participants were recruited in the greater Boston area using online and newspaper advertisements.

3.2 Apparatus

The experiment was conducted in the MIT AgeLab "Aware Car", an instrumented Volvo XC 90 equipped with a customized data acquisition system designed for time synchronized measurement of vehicle, driver and environmental factors. Data capture was facilitated through a number of embedded sensing systems including a MEDAC System/3 instrumentation unit and NeuGraph software (NeuroDyne Medical Corporation, Cambridge, MA) for physiological measurement. In addition to capturing data, the system included functionality for manual and time based triggering that was used for the presentation of an auditory prompt / verbal response task.

Three levels of difficulty of a delayed digit recall task (n-back) were employed to present drivers with a low, moderate and high level of secondary cognitive workload (0-, 1-, & 2-back respectively). The presentation order of the levels was "ramped-up" to systematically increase the demands on the driver. Further details on the development and presentation of the task can be found in Reimer [11].

3.3 Procedure

The experiment consisted of two counterbalanced assessments of identical presentations of the secondary task. One presentation took place in a driving simulator and the second, reported on here, in a field vehicle.

A modified lead II configuration was employed for EKG recording; the negative lead was placed just under the right clavical (collar bone), ground just under the left clavical, and the positive lead on the left side over the lower rib. The skin was cleaned with isopropyl alcohol and standard pre-gelled silver/silver chloride disposable electrodes (Vermed A10005, 7% chloride wet gel) were applied. Skin conductance level was measured utilizing a constant current configuration and non-polarizing, low impedance gold plated electrodes that allowed electrodermal recording without the use of conductive gel. Sensors were placed on the underside of the outer flange of the middle fingers of the non-dominant hand and held in place with medical grade paper tape.

During the field portion of the experiment, a research associate seated in the back of the vehicle observed the participant for signs of fatigue and other unsafe driving behaviors that could compromise safety. The research associate also operated the data collection equipment and provided driving direction. After leaving MIT, participants drove for approximately 30 minutes before stopping for a short break. Following the break, the experimental section of the protocol began. Instructions and other components of the experiment were prerecorded and played automatically over

the vehicle sound system. The eight segments of the experiment and the time allotted to each appear in Table 1. The posted speed limit was 65 mph (104.60 km/h).

Table 1. Experimental Protocol Overview

Segment	Duration (min:sec)	notes
Baseline	10:00	Minutes 7:22 – 9:30 used for the analysis period.
0-back training	1:04	Instructions and one 10 digit practice trial
0-back test	2:08	Four 10 digit evaluation trials
1-back training	2:08	Two repetitions of instructions and a 10 digit practice trial
1-back test	2:08	Four 10 digit evaluation trials
2-back training	3:12	Three repetitions of instructions and a 10 digit practice trial
2-back test	2:08	Four 10 digit evaluation trials
Recovery	7:00	Minutes 5:00 – 7:08 used for the analysis period.

3.4 Data Analysis

Due to recording issues, one participant's heart rate data and a second participant's skin conductance data were unavailable. A variety of algorithms were applied to the raw EKG and skin conductance data to remove noise and identify heart rate. Heart beats were detected using EKG Wave Editor release 1.8 (NeuroDyne Medical Corporation, Cambridge, MA), a software package that identifies R-wave peaks in the raw EKG signal and provides editing functionality. Processed records were reviewed by trained research associates and skipped and double beats were edited to provide a normalized heart rate record following general guidelines recommended for heart rate variability analysis [10]. To ensure accuracy and consistency, a second review of all records was performed by the second author.

Skin conductance is a "smooth" physiological signal characterized by relatively slow changing tonic levels and phasic sign wave like peaks associated with discrete arousal events. Artifacts associated with deformation of the sensor skin interface due to movement of the fingers or contact with the steering wheel typically appear as abrupt signal changes that are easily visualized. Artifact removal was carried out by first filtering skin conductance data using a wavelet decomposition with a *coif5* mother wavelet decomposed to level 3 using MatLab [9]. The signal was reconstructed using the level 3 approximation and detail coefficients using a similar *coif5* mother wavelet. This multi-level decomposition of the raw SCL signal into successive approximation and detail coefficients allows the breakdown of the SCL data into various lower resolution components of the original signal. This procedure allows the removal of high frequency artifacts without compromising the low frequency information in the SCL data. The second author directed manual removal of remaining low frequency movement artifacts. During this review, two cases were classified as outliers and dropped from the analysis.

Statistical comparisons were computed with SPSS 16 using a repeated measures general linear model (GLM) procedure and a Greenhouse-Geisser correction for models that violated the assumption of sphericity. Gender was evaluated and later dropped from the analyses presented here after being established as a non-

significant factor ($p > .05$). Pairwise t-tests were computed for significant results with a least significant difference (LSD) adjustment for multiple comparisons. Means and standard deviations reported in the figures are computed across participants where data exist for all levels of the comparison.

4. RESULTS

4.1 Heart Rate

Mean and standard deviation values by period are listed in Table 2 and mean heart rate is displayed graphically in Figure 1. Period significantly impacted heart rate, $F(1.97, 47.21) = 30.95$, $p < .001$, in a manner consistent with the position that heart rate varies with the level of workload. Overall, heart rate increased by 3.1 beats per minute (bpm) from baseline to the low secondary workload level (0-back task), by an additional 4.5 bpm at the moderate workload level (1-back task) and finally by another 1.1 bpm at the high workload level (2-back task). This represents an 8.7 bpm increase in heart rate from baseline to the highest level of secondary task engagement. All pairwise comparisons across the workload periods, except the change between the moderate and high workload periods, are significantly different ($p < .01$). Five minutes following the highest workload period heart rate recovers to within 0.6 bpm of the baseline value ($p > .05$). As would be expected with the variations in how participants respond to the secondary cognitive task, the standard deviation of heart rate increases with each level of task demand.

Table 2. Summary of Physiological Response Measures

	Heart Rate (beats / min)	Skin Conductance (micromhos)
Baseline	75.4 (9.2)	13.7 (4.8)
0-Back	78.4 (10.7)	15.0 (5.5)
1-Back	82.9 (11.9)	15.3 (5.7)
2-Back	84.0 (12.8)	15.3 (5.4)
Recovery	76.0 (9.2)	14.6 (5.2)

Note: Means with the standard deviations in parenthesis.

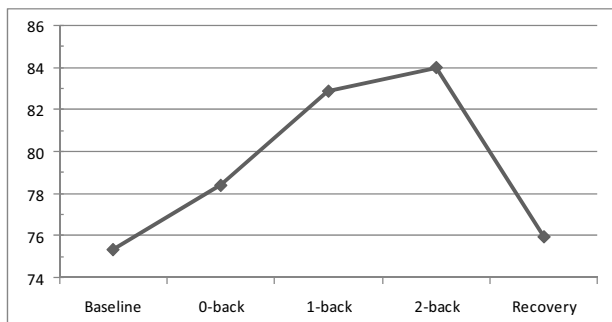


Figure 1. Mean heart rate (bpm) by period

Changes in mean heart rate from baseline to the highest level of task demand vary across subjects from -1.9 to 21.3 bpm. In two cases heart rate decreased. In one of these, the participant clearly had difficulty remaining engaged in the 2-back task, committing errors in 11 of 32 potential responses. This error rate was the largest exhibited in the sample. In the second case, the individual did not appear to be physiological responsive to the overt changes in demand, with heart rate varying only 1.9 bpm across the five periods. In five cases heart rate increased by less than three bpm and by 3 to 10 bpm in 8 cases. In the 10 remaining cases heart rate

increased by over 10 bpm between the baseline and high demand periods.

Comparing the data generated in this field study to results presented in Mehler et al. [8], incremental increases in heart rate across the demand levels appear strikingly similar (3.1, 4.5, and 1.1 vs. 3.1, 4.7, and 1.1 bpm respectively). Overall heart rate was higher in real vs. simulated driving, averaging 4.9, 4.8, 4.6 and 4.6 bpm higher over the first four periods respectively.

4.2 Skin Conductance

Mean skin conductance values by period are presented in Figure 2. As with heart rate, there is a significant effect of period, $F(2.32, 50.92) = 7.98$, $p = .001$. Continuing the pattern seen in the simulation study [8], the largest percentage increase in skin conductance occurs between the baseline and the initial / lowest secondary workload period (0-back). This mean change of 1.3 micromhos represents over 80% of the total rise in electrodermal activity observed across the three cognitive task periods. The incremental increase from the 0-back to the 1-back period was a more modest 0.3 micromhos and there was no additional increase in the mean during the 2-back. Pairwise comparisons show that baseline skin conductance differs from all other periods ($p < .005$ or less). However, skin conductance levels across the three tasks are not significantly different ($p < .05$) from each other. In contrast with heart rate, where there was essentially a return to baseline during the recovery period, skin conductance shows a more modest decrease. In fact, the change in mean values for skin conductance level during the 2-back and recovery period show a trend toward recovery but are not significantly different ($p = .07$).

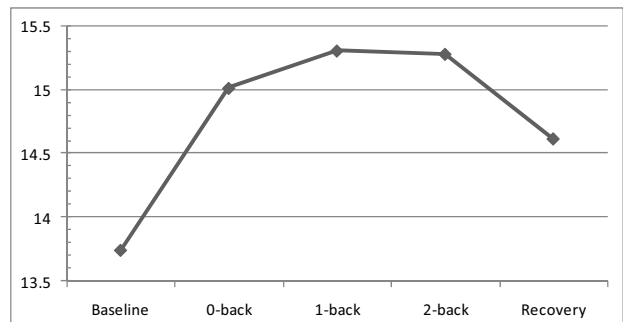


Figure 2. Mean skin conductance (micromhos) by period

5. CONCLUSION

Heart rate and skin conductance measures are presented as indicators of changes in driver workload associated with a secondary cognitive task. Consistent with Mehler et al. [8], heart rate appears to be a robust measure of incremental changes in real time workload associated with the secondary task. The current study replicates our simulation work by showing a remarkably consistent pattern of heart rate change in an on-road setting using the same protocol. Further, it extends this work by presenting data showing good recovery of heart rate to baseline levels following the cessation of the secondary load. Skin conductance data collected in the field also show a pattern similar to that seen in the simulation data. Skin conductance is quite sensitive to an initial increase in cognitive demand but appears to reach a ceiling effect more quickly. It is interesting to observe that the increase from baseline to the initial / low demand 0-back task was 2.4 micromhos in the simulator versus 1.3 on-road. One explanation

might be that the base level of arousal during actual driving is greater than in the simulator which would reduce the magnitude of the skin conductance level change if it tends to reach a ceiling level more readily. Skin conductance also differs from heart rate in that it appears to show a less rapid recovery following the completion of the secondary tasks. It may be that this slower recovery represents some remaining emotional engagement with the secondary tasks that is more evident in electrodermal activity than in heart rate.

The results presented here, in combination with task performance and vehicle performance data collected in this study [11], provide additional clarity as to the potential role of physiology in the detection of cognitive workload. Reimer [11] shows that drivers remained engaged in all levels the secondary tasks, while driving performance, characterized through lane keeping and forward velocity, only exhibits signs of deterioration at the highest level of secondary workload. A significant decrease in the standard deviation of gaze was also observed at this time. The physiological data presented here show that only a modest additional increase in heart rate occurred between the medium and high workload conditions, suggesting that drivers were nearing their available cognitive capacity for the combined load of driving and secondary tasks.

Aging and the corresponding health related declines in physical and cognitive function play an obvious role in reducing spare cognitive capacity [14]. These same factors may also potentially impact the sensitivity of real-time detection of workload through physiological measures. For example, various cardiac conditions and medications commonly used by older adults might reduce the effectiveness of using heart rate as a means of detecting driver workload by reducing beta-adrenergic reactivity.

The sequential presentation of the task demands from low to medium to high was carried out to allow a direct comparison with the earlier simulation [8] study. It is reasonable to question the extent to which the pattern of results obtained was influenced by presentation order and further investigation looking at the various load levels presented in a random order is clearly warranted. A consideration of presentation order effects and further assessment of age and health associated characteristics is currently underway. Continued exploration of these factors and other individual characteristics is important for determining the usefulness of physiological indices as a component of an algorithm for the real-time detection of driver workload.

6. ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of Ford Motor Company, the Santos Family Foundation and the United States Department of Transportation's Region I New England University Transportation Center at the Massachusetts Institute of Technology. We also wish to acknowledge the contributions of Jonathon Long, Vincent Lammers and Alea Mehler for their efforts in the collection and processing of data.

7. REFERENCES

- [1] Andreone, L., A. Amditis, E. Deregibus, S. Damiani, D. Morreale, and F. Bellotti. Beyond Context-Awareness: Driver-Vehicle-Environment Adaptively. From the COMUNICAR Project to the AIDE Concept. Proc. IFAC 16th World Congress, (2005).
- [2] Brookhuis, K.A. and D. De Waard, The use of psychophysiology to assess driver status. *Ergonomics*, Vol. 36, 9, (1993), 1099-1110.
- [3] Brookhuis, K.A. and D. De Waard, *Assessment of drivers' workload: Performance and subjective and physiological indexes*, in *Stress, Workload, and Fatigue*, P.A. Hancock and P.A. Desmond, Editors. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 321-333, 2001.
- [4] Brookhuis, K.A., G. De Vries, and D. De Waard, The effects of mobile telephoning on driving performance. *Accident Analysis and Prevention*, Vol. 24, 3, (1991), 309-316.
- [5] Broström, R., J. Engström, A. Agnvall, and G. Markkula. Towards the next generation intelligent driver information system (IDIS): The Volvo car interaction manager concept. Proc. 2006 ITS World Congress, (2006).
- [6] Engström, J., E. Johansson, and J. Östlund, Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F*, Vol. 8, 2, (2005), 97-120.
- [7] Lenneman, J.K., J.R. Shelley, and R.W. Backs. Deciphering psychological-physiological mappings while driving and performing a secondary memory task. Proc. Third International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, (2005), 493-498.
- [8] Mehler, B., B. Reimer, J.F. Coughlin, and J.A. Dusek. The impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers. Proc. Transportation Research Board of The National Academies, (2009).
- [9] Misiti, M., Y. Misiti, G. Oppenheim, and J. Poggi. *Wavelet Toolbox 4 User's Guide*. 2009 [cited 2009 May 29]; Available from: http://www.mathworks.com/access/helpdesk/help/pdf_doc/wavelet/wavelet_ug.pdf.
- [10] Mulder, L.J., Measurement and analysis methods of heart rate and respiration for use in applied environments. *Biological Psychology*, Vol. 34, 2-3, (1992), 205-236.
- [11] Reimer, B. Cognitive Task Complexity and the Impact on Drivers' Visual Tunneling. Proc. Transportation Research Board of The National Academies, (2009).
- [12] Reimer, B., J.F. Coughlin, and B. Mehler. Development of a Driver Aware Vehicle for Monitoring, Managing & Motivating Older Operator Behavior. Proc. ITS-America, (2009).
- [13] Saab Innovations. 2009 [cited 2009 May 27]; Available from: <http://www.saabsunited.com/2005/12/saab-innovations.html>.
- [14] Salthouse, T.A., The processing-speed theory of adult age differences in cognition. *Psychological Review*, Vol. 103, 3, (1996), 403-428.
- [15] Wilson, G.F., *Psychophysiological test methods and procedures*, in *Handbook of Human Factors Testing and Evaluation*, S.G. Charlton and T.G. O'Brien, Editors. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 127-156, 2002.
- [16] Wu, C. and Y. Liu, Queuing network modeling of driver workload and performance. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 8, 3, (2007), 528-537.