

OPERATOR FATIGUE ESTIMATION USING HEART RATE MEASURES

R. Hefner¹, D. Edwards¹, C. Heinze², D. Sommer², M. Golz², B. Sirois³, U. Trutschel³

¹ Machine Research, Caterpillar Inc.
Peoria, Illinois, USA

² Faculty of Computer Science, University of Applied Science
Schmalkalden, Germany

³ Circadian Technologies Inc.
Stoneham, Massachusetts, USA
Email: utrutschel@circadian.com

Summary: This study was designed to evaluate the viability of utilizing Tachograms for estimating fatigue in industrial and transportation applications. To explore this possibility Tachograms were recorded continuously and several heart rate measures were calculated and correlated with other well established fatigue measures. It was anticipated that changes in operator fatigue during a night time study could be depicted during three different conditions. In the first condition, a 40-minute monotonous driving task, Karolinska Sleepiness Scale (KSS), Variation of Lane Deviation (VLD), number of Micro-Sleep Events (MSE), numbers of accidents, and the PERCLOS score were collected as subjective and objective fatigue measures. In the second condition, a 10-minute Compensatory Tracking Task (CTT), the Mean Distance (MD) of a moving disk to a given target, the Standard Deviation of the Distance (SDD), as well as the Mean Velocity (MV) of the disk and the Standard Deviation of the Velocity (SDV) over the test duration were used as fatigue measures. In the third condition, a 5-minute Psychomotoric Vigilance Test (PVT), the Mean Response Times (MRT), the Standard Deviation of the Response Times (SDRT), the Mean of the inverse of the Slowest 10% of Response Times (MS10% 1/RT), and the number of lapses were used as fatigue measures. Correlations between heart rate and fatigue measures were calculated and classified using experimental results of one volunteer, who completed two nighttime episodes in a real-car lab following a partial sleep deprivation design. Results show strong correlations between heart rate variability (HRV) measures and multiple fatigue measures.

INTRODUCTION

Objectively assessing operator fatigue is a difficult and time consuming task. Nevertheless, worker fatigue is a major concern for any operation that requires sustained employee alertness and performance. This is especially apparent in industries where 24/7 shift work schedules are utilized, such as in mining, transportation and emergency services. A fatigued workforce creates an increased risk for human error, which in many instances can have significant consequences. Numerous methods and technologies for objective fatigue monitoring have been proposed, based either on measures of operator actions (EEG recordings, eye activity, such as PERCLOS, and head movements), or based on equipment parameters (such as steering activity or lane deviation). In the 1970's and 80's heart rate measures were investigated regarding their ability to

indicate fatigue. In 1985 Bishop et al. found that during extended periods of low workload under monotonous conditions, Heart Rate (HR) in beats per minute decreased while Heart Rate Variability (HRV) increased. In subsequent research studies, the correlations between HR measures and operator fatigue were ambiguous. These results were attributed to the complex nature and sensitivity of HR measures. Resting HR varies widely by individual, and during mental and physiological stress HR can increase as much as up to three-fold. In healthy individuals, the HR at any time represents the net effect of the parasympathetic (vagus) nerves, which slow the HR, and the sympathetic nerves which accelerate the HR. Depending on individual levels of mental and physical fitness, resting HR can vary between 50 and 100 beats per minute. This leads to substantial difficulties for the interpretation of HR measures when trying to assess operator drowsiness objectively. In addition, HR measures vary due to many mental and physical influences, and also due to insufficient control over the environment, the time of day, and the nature of the primary operator task. HR increases and HRV decreases during increasing mental workload showing exactly the opposite characteristics found for increasing fatigue. Other HR measures such as Low Frequency (LF) and High Frequency (HF) bands, as well as their ratio, are showing similar contradicting sensitivities to workload and fatigue. To add to this complexity, Coumel et al. (1995) pointed out that several HR measures (HR, HRV, pNN50, HF-band) show a strong circadian behavior under normal sleeping conditions. Therefore, a correlation to fatigue can be expected. As a consequence, the behavior of HR measures in general, and their ability to indicate fatigue in particular, can only be established under strict experimental conditions eliminating the influence of mental workload. In this experimental protocol of driving and testing (CTT and PVT), the mental workload and their influences were reduced to a minimum. In addition, how the data are analyzed is equally important. Many research studies calculate the means of HR measures of a group of subjects performing one task. This can lead to contradicting results because of individual differences in physical and mental work capacity, fitness, chronotype, workload capacity and fatigue tolerance. Lee et al. (2003) proposed tracking the HR measures of one individual airplane pilot through several different phases of simulated flight. We believe this approach is useful for studying the complex interaction of workload, fatigue, and time of day, and that it will help to separate the different influences on the HR measures in general.

DATA COLLECTION

The study was conducted in the Department of Informatics, main research “Adaptive Biosignal Analysis”, at the University of Applied Sciences, Schmalkalden (Germany), and consisted of overnight driving simulation sessions (Fig. 1). Prior to the experiment, volunteers were trained on driving with the simulator and on different tests. Each volunteer was equipped with an activity monitor, and required to complete a sleep/wake log for at least 24 hours prior to the experiment. Volunteers arrived at the lab at 10 pm. After wire-up, checking of logs and activity monitors, experiment sessions started at 11:30 pm. Eight experimental sessions were conducted, each lasting one hour, and the last session finished at 8:30 am. Volunteers were given a 1-h break at 3:30 am. Each session included: A 40-minute driving session, 10-minute Compensatory Tracking Task (CTT), and a 5-minute Psychomotoric Vigilance Test (PVT). Karolinska Sleepiness Scale (KSS) and brief alertness assessments were also performed at regular intervals during the driving task. The data presented is from one participant who was tested for two nights in the real-car driving simulator lab.

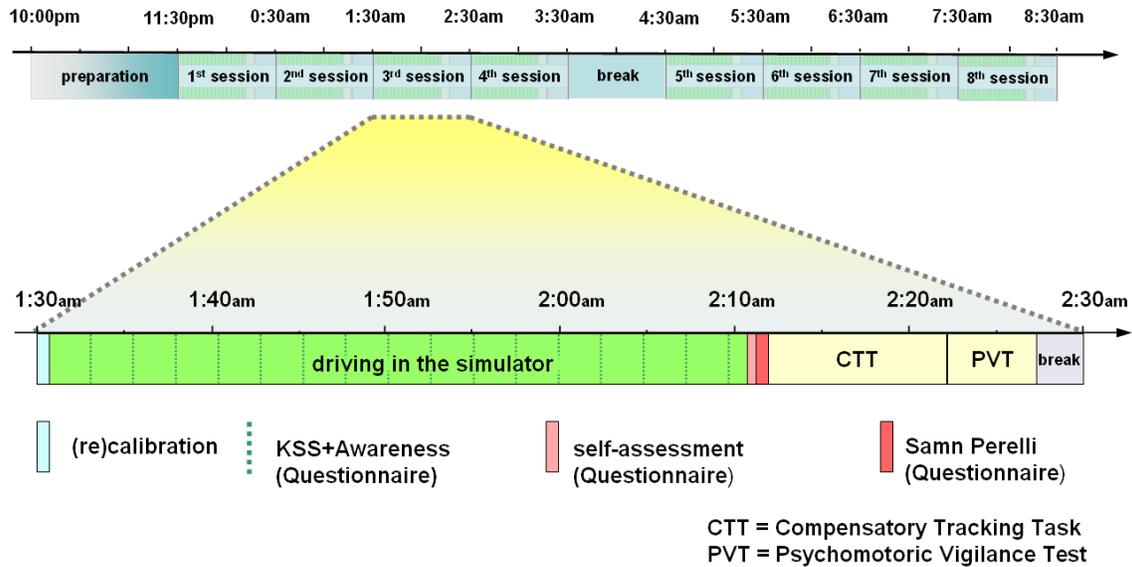


Figure1. Study design; HR was recorded continuously through the night, but the HR-measures were calculated for each driving, CTT, and PVT session and correlated with the corresponding fatigue measures.

DATA ANALYSIS

A. Heart Rate Measures

Time-Domain: HR measures can be calculated directly from the ECG time series. The simplest time domain measures are the Mean and Standard Deviation of the RR intervals which we denote as “MRR” and “SDRR”. Another simple HR-measure was calculated scaling the RR time series by the well-known z-transform and counting the Zero Crossings (# ZCrossing).

Frequency Domain: The RR interval time series is an irregularly time-sampled signal. Therefore, the RR interval time series was interpolated and resampled before Fast Fourier Transform (FFT) was applied, and before the Total Power Spectral Density (TPSD) was calculated. The PSD was analyzed by calculating the relative powers for different frequency bands. The commonly used frequency bands are Very Low Frequency (VLF) ranging from 0.001 Hz - 0.04 Hz, Low Frequency (LF) from 0.04 Hz - 0.15 Hz, and High Frequency (HF) from 0.15 Hz - 0.4 Hz. Our frequency-domain parameters include the mean relative powers of VLF, LF, and HF bands, and the standard deviation of the VLF-band (SDVLF).

Phase Space: It is realistic to presume that RR time series also contains nonlinear properties because of the complex regulation of the heart. A simple method to capture this nonlinear characteristic is to go into the phase space and use the so-called Poincare plot. It is a graphical depiction of the correlation between consecutive RR beats. The standard deviations along the semi-minor and semi-major axis of the ellipse are denoted “SD1” and “SD2”, respectively. SD1 describes short-term variability and SD2 describes the long-term variability of the heart rhythm.

B. Fatigue Measures

Driving: During each 40-minute driving session the number of visually scored Microsleep Events (# MSE), the number of Accident Events (# AE), the Karolinska Sleepiness Scale (KSS) mean score, the Variation of Lane Deviation (VLD), and the mean PERCLOS (Percentage of Eyelid Closure recorded by a commercially available eyetracking system) were collected. The selected parameter set allows us to capture the different facets of driver fatigue. “# MSE” represents an objective, individual specific fatigue measure. “# AE” represents an additional objective fatigue measure, because it shows the consequences of not being able to stay awake and keep in the lane. The “KSS” is a standardized subjective self-rating of sleepiness and is an independent measure on a numeric scale between 1 and 9. VLD is an objective measure of driving performance. It is calculated as the standard deviation of the lateral position of the car in the lane. Finally, PERCLOS was promoted for many years as an excellent objective measure of fatigue. These five parameters made up the set of Fatigue measures utilized during driving.

CTT: Following the driving sessions, the CTT (Van Orden et al. 2000), a 10-minute performance test, had to be completed. The CTT is a visuomotoric tracking task involving a target positioned in the center of a two-dimensional display, and a moving disk. The task is to keep the disk in or near the center of the target annulus using a trackball whose movement supplied a restorative force to the disk in the direction of trackball motion. The position of the disk is recorded as a function of its previous position and velocity. The Mean Distance (MD) of the moving disk, the target itself, the Standard Deviation of the Distance (SDD), as well as the Mean Velocity (MV) of the disk and the Standard Deviation of the Velocity (SDV) are used as the CTT Fatigue measures set.

PVT: The PVT used in this study was the short, 5 minute version evaluated in Lamond et al. (2005). The analyzed PVT parameters were based on the Response Time (RT), defined as the latency between the appearance of the stimulus and the subject’s button-push response. The selected PVT fatigue measures set included the Mean Response Times (MRT), and the Standard Deviation of Response Time (SDRT), the Mean of the inverse of the Slowest 10% of Response Times (MS10% 1/RT), and the number of lapses (# Lapses).

RESULTS

During the driving simulation there were numerous accidents caused by fatigue. As anticipated, the important features of the accident characteristics were that HR showed the opposite behavior as they did for the indication of fatigue. Therefore, we expected a mixed picture in the study results during the driving condition. The subject’s fatigue increased during the night, but the number of accidents increased as well, triggering an apparent contradicting influence on the HR measures. This should not be the case during the CTT and PVT tests. They are designed to be monotonous, with low mental workload and without the danger of an accident. Here, a clearer relation between HR-measures and fatigue measures is expected.

An overview of the results is depicted in Figures 2a-c. Please note that the start times for the 8 sessions differ somewhat between the two nights. The sessions in the first night started and ended later due to some technical difficulties at the first night. All measures are z-scored for a

better comparison. During both nights, the subjects exhibited an extreme struggle while driving (Figure 2a). The worst times were between 5.30 and 7.30 (Session 6 in the first night, sessions 6, 7, and 8 in the second night). An even clearer picture is shown by the fatigue measures of the CTT and PVT. The fatigue was significantly higher in the second half of the night. Under all three conditions (driving, CTT, PVT) a time-on-task and modest time-of-day behavior regarding both fatigue measures and HR measures was clearly recognizable. In general, the multiple HR measures correlated very well with the multiple fatigue measures for all three conditions. This is a very encouraging first result. The one hour break in the middle of the night reduced the fatigue levels during the fifth driving session (Figure 2a), but not for the subsequent CTT and PVT session (Figure 2b, 2c). The results of the last session in the first night are showing an improvement in fatigue, reflecting the influence of the circadian rhythm in the morning.

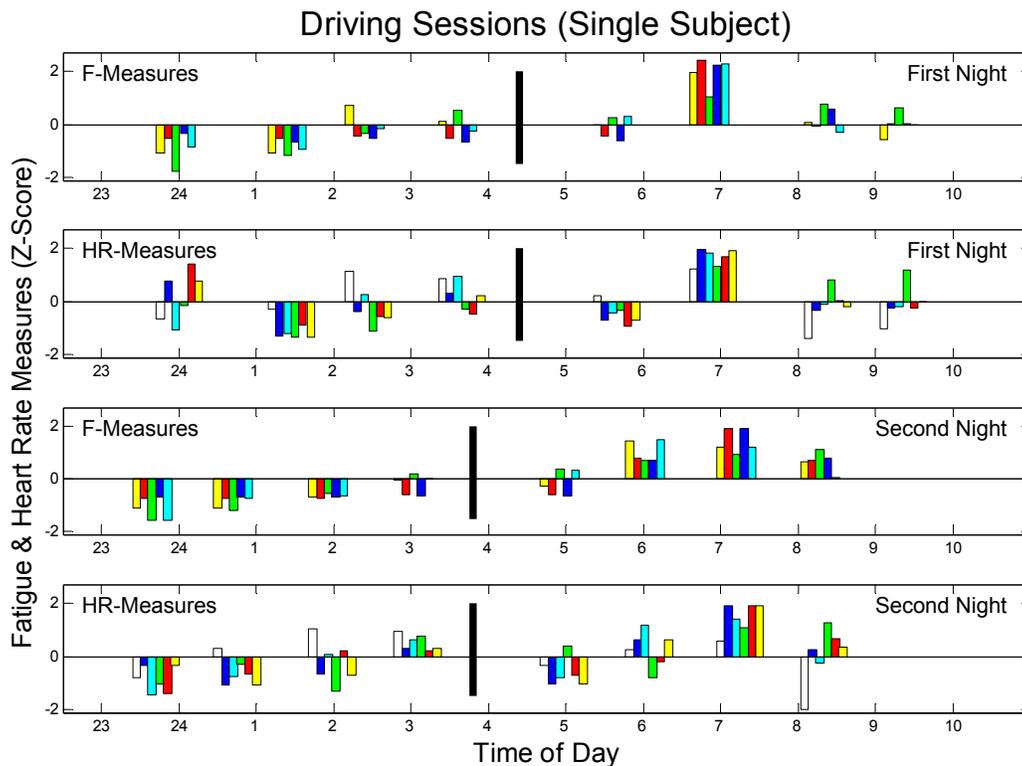


Figure 2a. Fatigue measures & HR measures during eight consecutive driving sessions. Panel 1 & 3 depict several common fatigue measures (from left to right): # MSE (yellow), # AE (red), KSS (green), VDL (blue), PERCLOS (cyan), and Panel 2 & 4 depict the several HR measures (from left to right): MRR (white), SDRR (blue), TPSD (cyan), VLF (green), SDVLF (red), SD2 (yellow). Vertical black bar in the middle depicts the one hour break. All measures were z-transformed.

Please note, that, in the different Figures 2, not all and not the same HR measures are depicted. One reason is for clarity, but another reason is that certain HR measures are not correlating well with some of the common fatigue measures (F measures). For example, HR measure TPSD is depicted in Figure 2a and 2b, but not in Figure 2c. For the PVT sessions, the F measure “MS10% 1/RT” showed an inverse behavior similar to the HR measure “#ZCrossings”, which was added to Figure 2c.

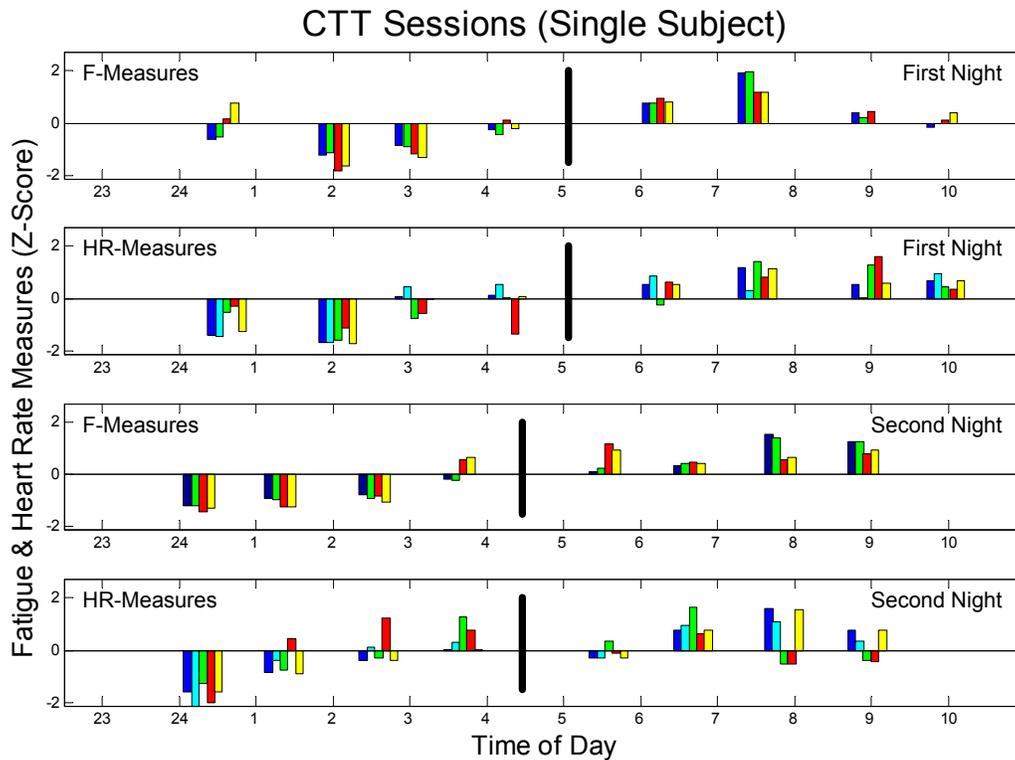


Figure 2b. F measures and HR measures during eight consecutive CTT sessions. Panel 1 & 3 depict the CTT F measures (from left to right): MD (blue), SDD (green), MV (red), SDV (yellow), and Panel 2 & 4 depict the HR measures (from left to right): SDRR (blue), TPSD (cyan), VLF (green), SDVLF (red), SD2 (yellow). Vertical black bar in the middle depicts the one hour break.

Figure 3 shows the important correlations between HR measures and F measures for the three different conditions. The results are somewhat surprising. The HR (MRR) itself shows no significant correlation to any of the recorded F measures, even under the very controlled conditions of the CTT and PVT. The same was true for “pNN50”. But other HR measures, such as SDRR, #ZCrossings, and SD2, display strong correlations and can be considered as strong indicators of fatigue.

Surprisingly, some HR-measures were only useful to indicate fatigue under certain conditions. For example, during the driving sessions TPSD was a very strong indicator, but it failed for CTT and PVT. The LF-band correlated only with the subjective Fatigue measure KSS. The HF-band shows a negative correlation on a significant level with KSS for driving, and MV and SDV for CTT. This was expected because of the antagonist relation between LF and HF bands, representing the sympatho-vagal balance. The vagal influence should be predominant over the sympathetic drive at night. This was observed for the HR-measures SDRR, SD2, and #ZCrossings, but not for the HF-band in all three conditions. Strong correlations occurred between SDVLF and driving performance (# AE, and VLD). For the PVT, the fatigue measure “# Lapses” has some potential but never showed significant correlation with any HR measure. Overall, several HR-measures can be used as reliable fatigue indicator.

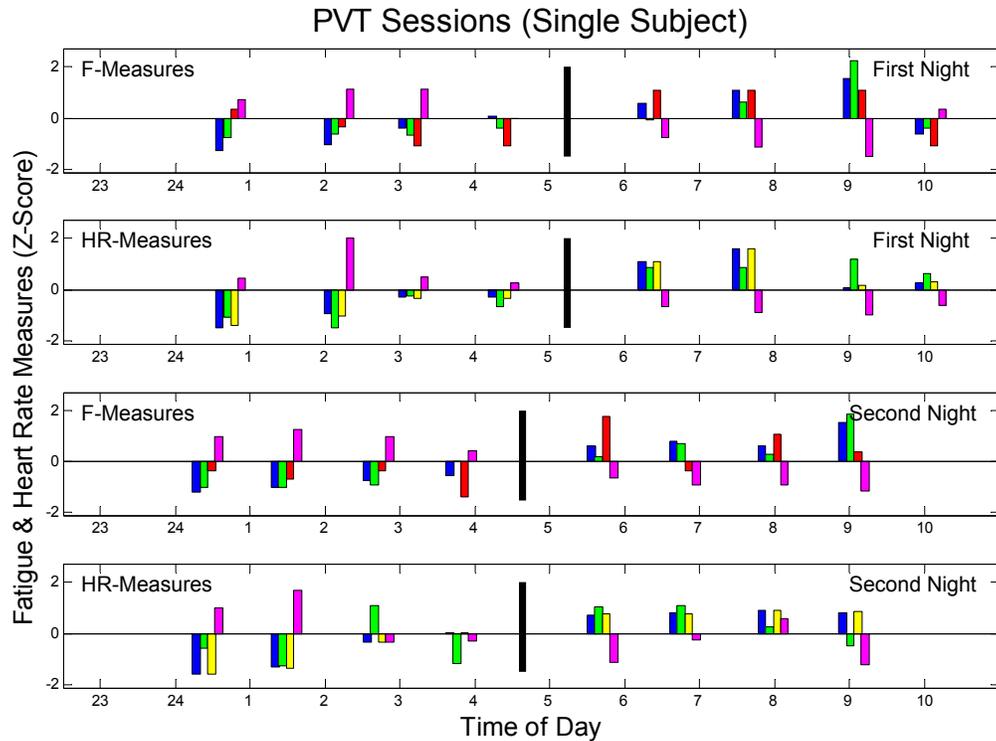
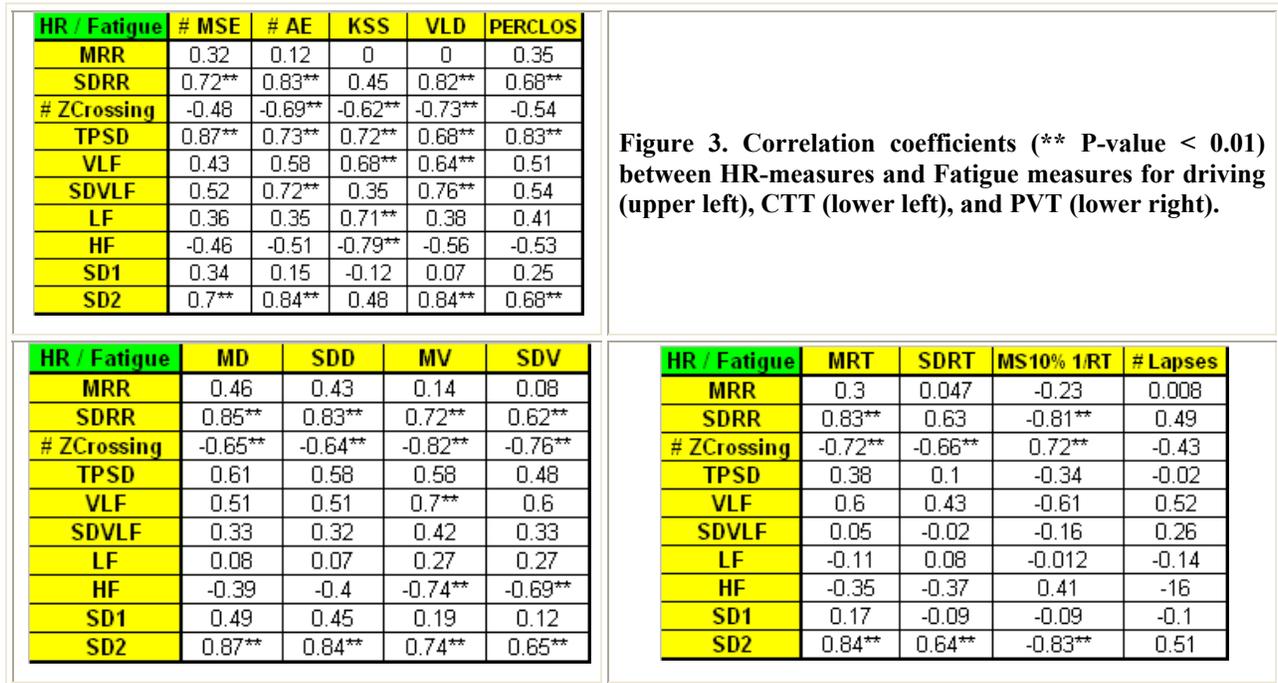


Figure 2c. F and HR measures during eight consecutive PVT sessions. Panel 1 & 3 depict the F measures of the PVT (from left to right): MRT (blue), SDRT (green), # Lapses (red), MS10% 1/RT (magenta), and Panel 2 & 4 depict the HR measures (from left to right): SDRR (blue), VLF (green), SD2 (yellow), and # ZCrossings (magenta). Vertical black bar in the middle depicts the one hour break.



CONCLUSIONS

Under the specific task conditions of low mental workload such as in overnight driving, CTT and PVT sessions, several HR measures were found to be promising indicators of drowsiness and fatigue. The previous findings by Bishop et al. (1985) were partly confirmed. HRV measures in general (but not the HR itself) are good indicators of fatigue. Strong correlations between fatigue measures and HR measures are observed in SDRR, #ZCrossings, and SD2. The selected path to correlate multiple HR measures with multiple fatigue measures based on an individual subject appear to be successful for establishing the fatigue estimation capabilities of HR measures. Nevertheless, to confirm these results, and to generalize our findings will require data collection and analysis of additional subjects.

REFERENCES

- Bishop, H., Madnick, B., Walter, R., Sussman, E.D. (1985). Potential driver attention monitoring system development. (Final Report No. DOT-TSC-NHTSA-85-1).
- Coumel, P., Maison-Blanche, P., Catuli, D., Heart rate and heart rate variability, in Heart rate variability, Malik, M., Camm, A.J., Futura Publishing Company, Inc. 1995.
- Lal, S. Craig, A. Driver fatigue: electroencephalography and psychological assessment. *Psychophysiology*, vol. 39, pp. 313–321, 2002.
- Lamond, N., Dawson, D., and Roach, G.D. Fatigue Assessment in the Field: Validation of a Hand-Held Electronic Psychomotor Vigilance Task, *Aviation, Space, and Environmental Medicine*, vol. 76, 2005
- Lee, Y-H., Liu, B-S., Inflight Workload Assessment: Comparison of Subjective and Physiological Measurements, *Aviation, Space, and Environmental Medicine*, 74, 1078-1084, (2003).
- Van Orden, K.F., Jung, T.P., Makeig, S., Combined eye activity measures accurately estimate changes in sustained visual task performance, *Biological Psychology* 52 (2000) 221–240.