Operator Fatigue Estimation Using Heart Rate Measures

C. Heinze[1], U. Trutschel[2], T. Schnupp[1], D. Sommer[1], A. Schenka[1], J. Krajewski[3], M. Golz[1]

[1]Faculty of Computer Science, University of Applied Sciences Schmalkalden, Schmalkalden, Germany

[2]Circadian Technologies Inc., Stoneham, Massachusetts USA

[3]Institute of Work and Organizational Psychology, University of Wuppertal, Wuppertal, Germany

Abstract—**The growing number of fatigue related accidents in recent years has become a serious concern. Accidents caused by fatigue in transportation and in mining operations involving heavy equipment can lead to substantial damage and loss of human life. Preventing such fatigue related accidents is highly desirable, but requires techniques for continuously estimating and predicting the operator's alertness state. This paper proposes ECG-based operator fatigue estimation. For this aim, ECG was recorded continuously, and several heart rate measures were calculated and correlated with other well established fatigue labels. As a result, changes in operator's fatigue during a night time study could be depicted during three different conditions. In the first condition, subjective and objective fatigue measures were collected during a 40-minute monotonous driving task. In the second and third condition, a 10-minute Compensatory Tracking Task (CTT) and a 5-minute Psychomotoric Vigilance Test (PVT), respectively, delivered a set of additional objective fatigue measures. Correlations between heart rate and fatigue measures were calculated, using experimental results of two volunteers, who each completed two nights in a real-car lab following a partial sleep deprivation design. The subjects were going through all three conditions (driving, CTT and PVT) eight times during the course of one night.**

Keywords—**Fatigue, Heart Rate Variability, Overnight Driving, CTT, PVT**

I. INTRODUCTION

Objectively assessing operator fatigue is a difficult task. Nevertheless, worker fatigue is a concern for any operation that requires an employees' sustained alert performance. This is especially apparent where 24/7 shift work schedules are standard, e. g. as in mining, transportation and emergency services.

In the 70'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 has been shown to decrease while heart rate variability (HRV) has been shown to increase. These effects and their relationship to fatigue have been demonstrated in an 'on-the-road' driving task [O'Hanlon & Kelly] and simulator based environment [Wierwille & Muto].

During mental and physiological stress HR can increase up to three-fold. In a healthy individual, the HR at any time represents the net effect of the parasympathetic nerves which slow down the HR and the sympathetic nerves on the other hand which accelerate the HR. Resting HR varies widely per individual (between 50 and 100 beats per minute), which leads to substantial difficulties for the interpretation of HR-measures in order to assess operator fatigue objectively. In addition, HR-measures may vary because of many mental and physical influences and insufficient control over the environment. For example, HR increases and HRV decreases during increasing mental workload, whereas the exact opposite characteristics can be found for increasing fatigue. To add to this complexity, [Coumel et al] pointed out that several HR-measures show a strong circadian behavior; therefore a correlation to fatigue can be expected. As a consequence, the behavior of HR-measures in their ability to indicate fatigue can only be established under strict experimental conditions, eliminating the influence of mental workload.

In our situation, the experimental protocol of driving and testing (CTT and PVT) reduced the mental workload and their influences to a minimum. In addition, we assessed the HR-measures of our subjects individually, because in our opinion calculating the means of HR-measures over a group of subjects can lead to contradicting results.

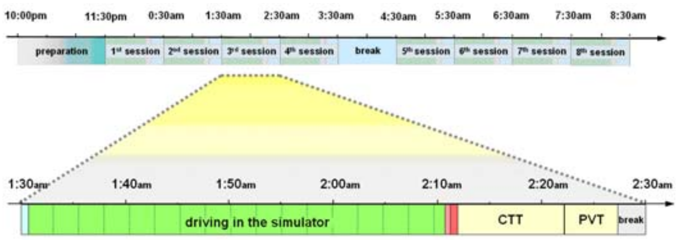

Figure 1: Design of one overnight experiment, divided in eight sessions. ECG was recorded continuously through the experiment. HR-measures were calculated for each driving, CTT and PVT session and correlated with the corresponding fatigue measures.

II. EXPERIMENTS

Our study consisted of two nights of driving simulation (in a real-car driving simulator lab) per participant. The design of one overnight experiment is shown in fig. 1. For each participant, there was a gap of at least one week between the two experiments. At the start of the experiment (11:30 pm), each

O. Dössel and W.C. Schlegel (Eds.): WC 2009, IFMBE Proceedings 25/IV, pp. 930–933, 2009.
www.springerlink.com

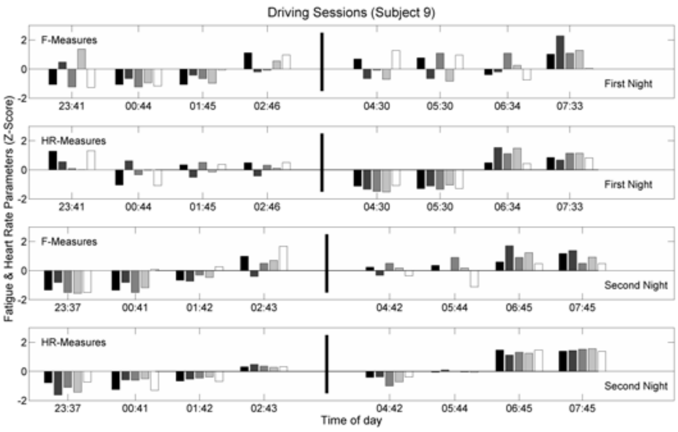

Figure 2a: Two nights of eight consecutive *driving sessions* for subject 9. Fatigue measures (black to white): MSE, ACC, KSS, VLD, PCLOS. HR-measures (black to white): SD RR, pNN50, M TPow, SD1, SD2.

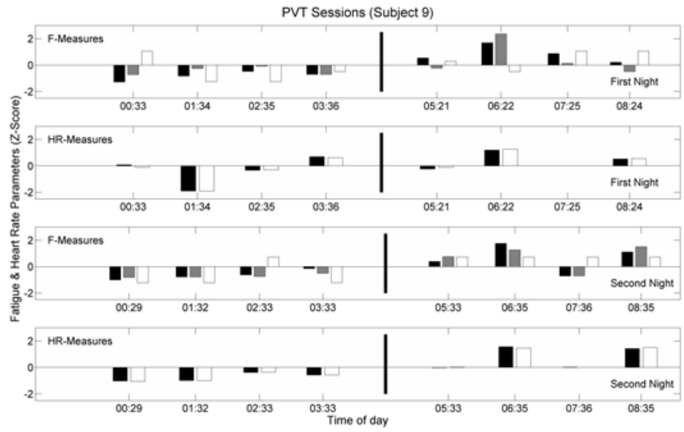

Figure 2c: Two nights of eight consecutive *PVT sessions* for subject 9. Fatigue measures (black to white): M Latc, SD Latc, #Lps. HR-measures (black to white): SD RR, SD2.

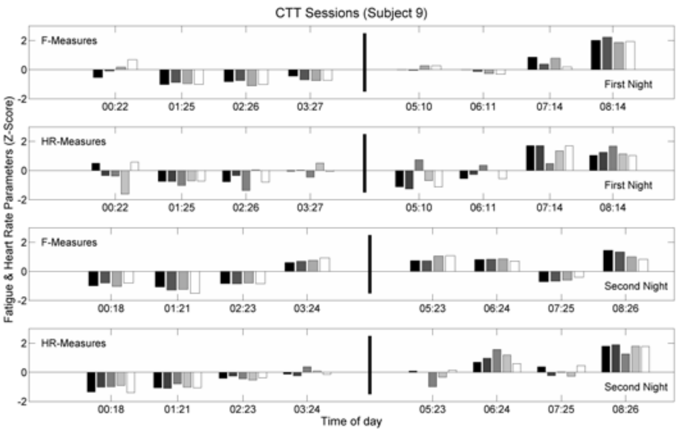

Figure 2b: Two nights of eight consecutive *CTT sessions* for subject 9. Fatigue measures (black to white): M Dist, SD Dist, M Velc, SD Velc. HR-measures (black to white): SD RR, M TPow, SD TPow, SD1, SD2.

participant had an assured uninterrupted time-since-sleep of at least 18 hours. The entire study included 12 participants; the data presented here is from participants no. 9 and 11.

III. Methods

A. Heart Rate Measures

Time Domain HRV is defined as the consecutive beat-to-beat (RR) time intervals of the ECG. The simplest time domain measures are the *mean and standard deviation of this RR time series*, which we denote as 'M RR' and 'SD RR'. All adjacent *RR-intervals differing more than 50 msec* were counted and normalized by the length of the RR time series, resulting in the HR-measure 'pNN50'. Another simple measure was calculated by z-transforming the RR time series and counting the *zero-crossings* (denoted 'ZCross').

Frequency Domain The RR time series is an irregularly sampled signal, therefore it was interpolated before Fast Fourier Transform (FFT) was applied. The power spectral density was analyzed by calculating the relative powers for different frequency bands. 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 measures include the means and standard deviations of total power, VLF-, LF- and HF-bands (denoted 'M TPow', 'SD TPow', 'M VLF', 'SD VLF', 'M LF', 'SD LF', 'M HF', 'SD HF', respectively).

Phase Space It is realistic to presume that RR time series also contain nonlinear properties, because of the complex regulation of the heart. A simple method to capture this nonlinear characteristic uses the *Poincaré plot* in phase space, which is a graphical depiction of the correlation between consecutive RR beats in a shape of an ellipse. The standard deviations along the semi-minor and semi-major axis of the ellipse are denoted 'SD1' and 'SD2', respectively. The first describes *short-term*, the latter the *long-term variability* of the heart rhythm.

B. Fatigue Measures

Driving Session During driving, *micro-sleep events* (MSE) were visually scored. The occurrence of MSE strongly varies from person to person; typical visible signs are prolonged eyelid closures, head nodding, spontaneous pupil contractions and stare gaze. MSE are an objective, individual-specific fatigue measure. The number of each MSE during one driving session is weighted by its severity; the weighted sum constitutes the fatigue measure 'MSE'. *Accident events* (AE) represent an additional objective fatigue measure, because they show the consequences of not being able to stay awake and keep the lane. AE are also weighted and summed up to a measure denoted as 'ACC'. The *Karolinska sleepiness scale* (a standardized and subjective measure of fatigue on a numeric scale between 1 and 10) is polled in regular intervals during driving; its mean score constitutes the measure 'KSS'. The *variation of lane deviation* is recorded as the distance from the lane's center and is an objective measure of driving performance; its standard deviation constitutes the measure 'VLD'. Finally, *PERCLOS* (PERcentage of eyelid CLOsure, recorded by commercial eye-tracking system) was promoted for many years as an excellent fatigue measure; its mean score constitutes the measure 'PCLOS'. These five parameters are making up the set of fatigue measures during driving.

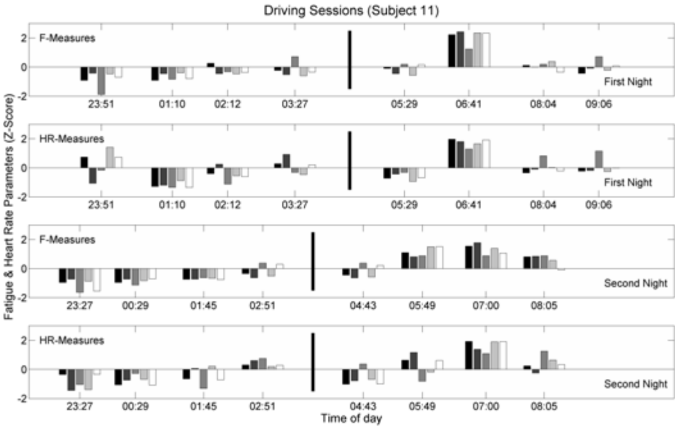

Figure 3a: Two nights of eight consecutive *driving sessions* for subject 11. Fatigue measures (black to white): MSE, ACC, KSS, VLD, PCLOS. HR-measures (black to white): SD RR, M TPow, M VLF, SD VLF, SD2.

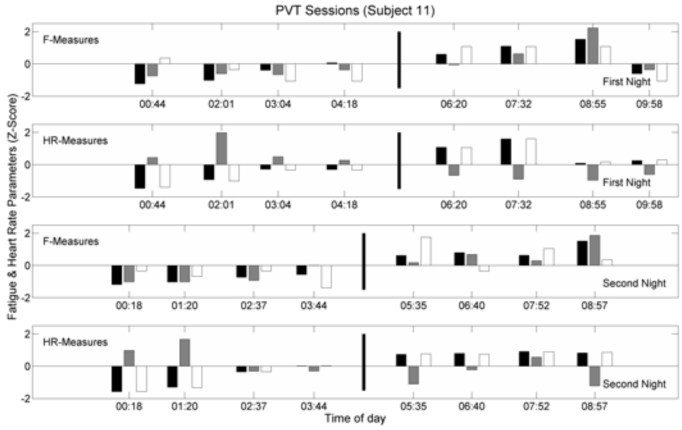

Figure 3c: Two nights of eight consecutive *PVT sessions* for subject 11. Fatigue measures (black to white): M Latc, SD Latc, #Lps. HR-measures (black to white): SD RR, ZCross (negative correlation), SD2.

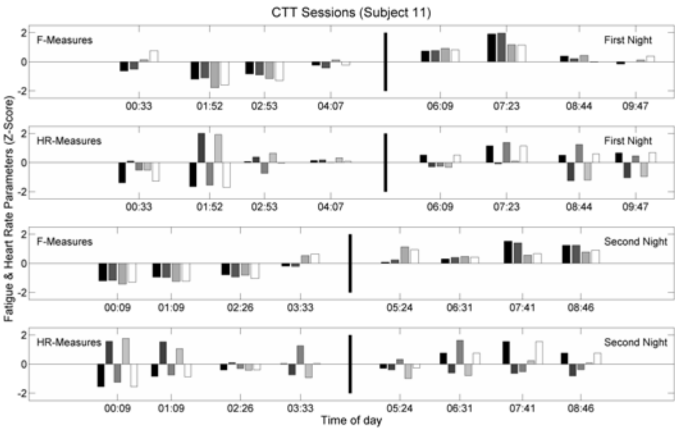

Figure 3b: Two nights of eight consecutive *CTT sessions* for subject 11. Fatigue measures (black to white): M Dist, SD Dist, M Velc, SD Velc. HR-measures (black to white): SD RR, ZCross (negative correlation), M VLF, M HF (negative correlation), SD2.

CTT Session A CTT performance test – described by [Van Orden et al] – had to be completed after driving. The CTT is a two-dimensional visual compensatory tracking task involving a target positioned in the center of a 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 and the standard deviation of the *distance* of the moving disk to the target (denoted 'M Dist' and 'SD Dist') as well as the mean and standard deviation of the *velocity* of the disk (denoted 'M Velc' and 'SD Velc') make up the fatigue measures during the CTT.

PVT Session The PVT is a simple stimulus-reaction test, where reaction times are interpreted as a performance measure. The short 5-minute version used in our study was evaluated in [Lamond et al]. *Latency* is defined as the time between the appearance of the stimulus and the subject's button-push response. As a measure set, we selected the mean and standard deviation of latenties (denoted 'M Latc' and 'SD Latc') and the number of *lapses* (absence of response, denoted '#Lps').

IV. RESULTS

Figures 2a-c and 3a-c show comparisons between fatigue and HR-measures for subjects 9 and 11, respectively. Please note that the start times for the eight sessions differ somewhat between two respective experiments. All measures are z-scored to adjust for scaling differences.

In general, the progression of the fatigue measures under all three conditions and for both subjects shows a real struggle against fatigue during the course of the experiments, also there is a clear time-on-task behavior regarding fatigue. Multiple HR-measures from time domain, frequency domain and phase space are correlating with multiple fatigue measures. However, individual differences in the results are considerable.

For subject 9, the worst time of driving is between 6:30 and 8:30 am (7th and 8th session), where the fatigue measures reach top level (fig. 2a). An even clearer picture is provided by the CTT condition: Fatigue measures consistently increase during the course of the experiment and correlate with selected HR-measures; this is especially true for the 2nd night (fig. 2b). The PVT condition gives a different assessment: the most fatigued time is around 6:30 am (6th session), which is quite earlier than in driving and CTT (fig. 2c).

Subject 11 is most fatigued in driving between 6:30 and 7:30 am (6th session in the 1st night, 7th session in the 2nd night, fig. 3a). The CTT condition is consistent with this picture, where fatigue levels are highest during the 6th session/1st night and 7th session/2nd night (fig. 3b). The PVT condition shows a somewhat delayed effect: the highest fatigue is reached around 9:00 am, during the 7th session/1st night and 8th session/2nd night (fig. 3c). The results of subject 11 display a remarkable influence of the circadian rhythm: the timing of both experiments had a considerable misalignment of 30-45 minutes, yet the maximum of fatigue was reached at approximately the same time. Also, towards the end of the experiment the fatigue levels are decreasing. This points out that not only time-on-task, but also time-of-day effects have a significant influence on fatigue progression.

Between the 4th and 5th session, the subjects had an extensive break of one hour. Considering time-on-task behaviour, both subjects should partially recover from their previous fatigue. However, judging by the fatigue levels in all three con-

Table 1: Correlation coefficients between all HR measures (rows) and fatigue measures (columns). Significant correlations (P< 0.01) are marked gray.

		HR-measures vs Driving fatigue					HR-measures vs CTT fatigue				HR-measures vs PVT fatigue		
		MSE	ACC	KSS	VLD	PCLOS	M Dist	SD Dist	M Velc	SD Velc	M Latc	SD Latc	#Lps
Subject 9	M RR	0,5788	0,5495	0,6290	0,6213	0,2099	0,5466	0,4443	0,3778	0,2831	0,3764	0,3966	0,1104
	SD RR	0,3378	0,7910	0,3399	0,8591	0,0010	0,7463	0,7043	0,7177	0,6430	0,7450	0,6486	0,6073
	pNN50	0,2319	0,6472	0,3997	0,7204	-0,1009	0,6883	0,5789	0,5217	0,3908	0,0393	0,1868	0,3153
	ZCross	-0,3131	-0,6874	-0,3157	-0,7530	0,0303	0,0811	0,0661	-0,0616	-0,0701	-0,6044	-0,5049	-0,4518
	M TPow	0,2918	0,7508	0,3856	0,7009	0,0299	0,8011	0,7358	0,6978	0,5852	0,1855	0,2215	0,4896
	SD TPow	0,5200	0,5902	0,3586	0,4577	0,4789	0,7973	0,7771	0,7575	0,6959	0,1976	0,2701	0,0747
	M VLF	0,6321	0,6107	0,5836	0,4142	0,2530	0,3302	0,2456	0,2718	0,1962	0,4484	0,3786	0,1344
	SD VLF	0,3254	0,0291	0,3100	0,2340	0,2380	0,5437	0,4180	0,4696	0,3001	0,1080	0,2720	0,1093
	M LF	-0,2978	-0,7986	-0,2689	-0,7225	0,2347	-0,5459	-0,4651	-0,4014	-0,3362	-0,3064	-0,2878	-0,5148
	SD LF	0,6126	0,2198	0,6543	0,4246	0,4047	0,1156	0,0431	0,0533	-0,0378	0,4102	0,4298	0,2811
	M HF	-0,1395	0,2046	-0,1931	0,2643	-0,2008	0,2601	0,2489	0,1460	0,1471	-0,2228	-0,1672	0,2268
	SD HF	0,3125	0,3540	0,3891	0,2334	0,2032	0,2858	0,2086	0,2001	0,0733	0,3462	0,1130	0,4757
	SD1	0,3245	0,7387	0,4715	0,7222	0,0152	0,7729	0,6710	0,5926	0,4534	0,0436	0,1375	0,3611
	SD2	0,3383	0,7861	0,3336	0,8565	0,0012	0,7301	0,6909	0,7095	0,6397	0,7766	0,6721	0,6001
Subject 11	M HR	0,2146	0,0860	0,1226	0,0966	0,3447	0,4568	0,4259	0,1413	0,0823	0,3029	0,0470	0,0082
	SD RR	0,7818	0,8287	0,4640	0,7839	0,6996	0,8586	0,8311	0,7134	0,6237	0,8284	0,6247	0,4923
	pNN50	0,1091	0,0195	-0,1839	0,0491	0,0987	0,3549	0,3185	0,0422	-0,0330	0,0178	-0,2667	0,0042
	ZCross	-0,5931	-0,6891	-0,5794	-0,6353	-0,5607	-0,6471	-0,6383	-0,8162	-0,7628	-0,7174	-0,6616	-0,4328
	M TPow	0,8223	0,7084	0,7677	0,7383	0,8179	0,6123	0,5819	0,5812	0,4777	0,3822	0,0955	-0,0205
	SD TPow	0,4682	0,2558	0,5355	0,3744	0,5889	0,2076	0,2286	-0,0365	0,0196	0,2407	0,1699	0,0780
	M VLF	0,5224	0,5786	0,6247	0,5008	0,5295	0,5104	0,5144	0,6977	0,6095	0,6082	0,4253	0,5193
	SD VLF	0,6225	0,7314	0,3161	0,6694	0,5574	0,3276	0,3169	0,4189	0,3363	0,0559	-0,0234	0,2558
	M LF	0,3693	0,3515	0,6901	0,3987	0,4121	0,0778	0,0680	0,2674	0,2745	-0,1091	0,0779	-0,1383
	SD LF	-0,2049	-0,3000	-0,1206	-0,3242	-0,3506	0,0500	0,0325	0,1742	0,0504	0,1346	0,0300	0,0394
	M HF	-0,5063	-0,5169	-0,7460	-0,5192	-0,5361	-0,3884	-0,3966	-0,7398	-0,6890	-0,3430	-0,3709	-0,1620
	SD HF	0,1309	0,1270	-0,2137	0,0691	-0,0300	-0,1099	-0,0635	0,0522	-0,0277	0,1704	-0,0620	0,1968
	SD1	0,2479	0,1326	-0,0298	0,1750	0,2362	0,4902	0,4535	0,1926	0,1199	0,1644	-0,0919	-0,0987
	SD2	0,7734	0,8343	0,4821	0,7879	0,6974	0,8667	0,8403	0,7361	0,6507	0,8356	0,6416	0,5070

ditions, this is not the case, which hints at a considerable influence of time-of-day effects. Furthermore, for both subjects the PVT is surprisingly inconsistent with the other two conditions: where driving and CTT both point to their highest fatigue levels at approximately the same time-of-day, the PVT determines this fatigue maximum at another time.

Please note that in figures 2a-c and 3a-c not all and not the same HR-measures are displayed. This was done for clarity, because each fatigue measure correlates well only with some selected HR-measures (this is true for different conditions and different subjects). This can be seen in table 1, where a complete overview for correlations between all HR-measures and fatigue measures is given. Surprisingly, some HR-measures were only useful to indicate fatigue under certain conditions. For example, the HR-measure 'SD TPow' of subject 9 correlates consistently well with all CTT-fatigue measures, but not with any of the driving- or PVT-fatigue measures. Another example is 'M TPow' of subject 11, which indicates fatigue during the driving condition very well, but shows no correlation with any of the CTT- or PVT-fatigue measures. The HR-measures 'SD RR' and 'SD2' (additionally 'ZCross', only for subject 11) can be considered as the strongest indicators for fatigue, since they correlate with the most of the available fatigue measures. The most uncorrelated (and therefore unreliable) HR-measures are 'SD HF', 'M RR', 'SD LF' and 'SD VLF' (and additionally: 'M HF' and 'M VLF' for subject 9; 'SD TPow', 'SD1' and 'M LF' for subject 11).

V. Conclusions

Under specific task conditions of low mental workload such as overnight driving, several HR-measures are an excellent indicator of drowsiness. The previous findings by [Bishop et al] could be confirmed regarding his statement about HRV. Strong correlations between fatigue and HR-measures are observed in 'SD RR', 'ZCross' and the long term HR-measure 'SD2' obtained from the Poincaré map. The chosen way to rely on data from single subjects first for establishing the fatigue estimation capabilities of HR-measures turned out to be successful. Nevertheless, we will analyze the data of additional subjects over time. Based on our experience with other bio-signals such as EEG and EOG, substantial individual differences are expected.

References

[Bishop et al] Bishop H, Madnick B, Walter R & Sussman E D (1985): *Potential driver attention monitoring system development.* (Final Report No. DOT-TSC-NHTSA-85-1), National Highway Traffic Safety Administration, United States Department of Transportation. Washington, DC.

[Coumel et al] Coumel P, Maison-Blanche P & Catuli D (1995): *Heart rate and heart rate variability.* J Heart Rate Variability, Editors (Malik, M., Camm, A.J.), Futura Publishing Company, Inc. 1995

[Lamond et al] Lamond N, Dawson D & Roach G D (2005): *Fatigue Assessment in the Field: Validation of a Hand-Held Electronic Psychomotor Vigilance Task.* Aviation, Space and Environmental Medicine, vol. 76.

[O'Hanlon & Kelly] O'Hanlon J F & Kelly G R (1977): *Comparison of performance and physiological changes between drivers who perform well and poorly during prolonged vehicular operation.* In R.R. Mackie (Ed.) Vigilance: Theory, operational performance and physiological correlates. New York: Plenum Press.

[Van Orden et al] Van Orden K F, Jung T P & Makeig S (2000): *Combined eye activity measures accurately estimate changes in sustained visual task performance.* Biological Psychology, vol. 52, pp. 221-240.

[Wierwille & Muto] Wierwille W W & Muto W H (1981): *Significant changes in driver-vehicle response measures for extended duration simulated driving tasks.* Proceedings of the First European Annual Conference on Human Decision Making and Manual Control (pp. 298-314). Delft University of Technology, Delft, Netherlands.