Heart Rate Measures Reflect the Interaction of Low Mental Workload and Fatigue During Driving Simulation

Udo Trutschel  
Circadian  
2 Main Street  
Stoneham, MA 02180, USA  
utrutschel@circadian.com

Christian Heinze  
University of Applied Science  
Blechhammer 4-9  
98574 Schmalkalden, Germany  
c.heinze@fh-sm.de

Bill Sirois  
Circadian Technologies, Inc.  
2 Main Street  
Stoneham, MA 02180, USA  
bsirois@circadian.com

Martin Golz  
University of Applied Science  
Blechhammer 4-9  
98574 Schmalkalden, Germany  
golz@fh-sm.de

David Sommer  
University of Applied Science  
Blechhammer 4-9  
98574 Schmalkalden, Germany  
davegmx@sund.de

David Edwards  
Caterpillar  
300 Hamilton Blvd  
Peoria, IL 61602, USA  
edwards_david_j@cat.com

ABSTRACT
The objective of this study was to assess the monotonic mental workload under changing conditions of operator fatigue during a night time driver simulation study. Several cardiovascular measures were used in order to differentiate between driving and a continuous tracking task. From all of the standard cardiovascular measures, heart rate in beats per minute emerged as the most sensitive for workload discrimination. Heart rate was higher during driving than during the tracking task, pointing to a slightly higher demanding workload for the driving task. This result was stable over the course of the night and showed only a minimal fatigue influence. Heart rate variability in milliseconds, on the other hand, was on average higher for the continuous tracking task in comparison to the driving. This was especially the case for the sessions with high subjective sleepiness. It can thus be concluded that the fatigue state of the operator was more impaired during the tracking task than during driving.

Categories and Subject Descriptors

Keywords
Heart rate variability; Operator fatigue; Workload.

1. INTRODUCTION
It is well known that the heart seems to respond to almost every internal and external stimulus in our environment. Therefore, cardiovascular measures, particular heart rate (HR) and heart rate variability (HRV) are often used for assessing cognitive workload or operator fatigue in many fields of human factors, such as in aviation and automotive research [1-3]. Previous investigations focus either on workload or fatigue. But in nearly all working environments workload and human fatigue occur at the same time. For example, airline pilots and air traffic controllers are confronted with high mental workload and changing fatigue conditions from alert to sleepy. Truck drivers and railroad engineers experience low mental workload combined with fatiguing work environments during most of their duty periods. It is also well known that HR increases along with task demand. This seems to be valid even for small incremental steps of increasing cognitive workload [4]. On the other hand HRV decreases with strong increases in workload, but small incremental increases of workload cannot be detected. Motivated by the work of MEHLER ET AL., we reevaluated HR and HRV data from an overnight driving simulation scenario and a continuous tracking tasks (CTT). This driving simulation was originally designed without any other traffic in order to induce different levels of fatigue over the course of the night. The tracking task (CTT) was designed as assessment tool for measuring operator fatigue based on performance. Previously, we found that HRV increases with increasing driver fatigue for both tasks, but HR did not show any correlations with typical driving fatigue measures [5]. CHUA ET AL. reports a 67% correlation between HRV and the lapses of the psychomotor vigilance task (PVT), which strongly correlates with operator fatigue. From the study design depicted in fig. 1 it is clear that operator fatigue will increase during the course of the night. This is known as the time of day effect.

Thus, circadian influences do affect the operator performance. In addition, during each session there was an increase in operator fatigue based on time on task. On the other hand, the workload demand in all sessions is constant over time, but different for the driving and CTT. The goal of this current investigation was to utilize HR and HRV measures to compare the workload between driving in a simulator and a tracking task (CTT). With respect to the workload discrimination the following questions can be asked: Which cardiovascular measure is most sensitive for load and/or...
fatigue discrimination between the tasks, and which one is the better indicator for performance impairment during driving and CTT?

2. STUDY DESIGN

The overall study design is depicted in fig. 1. Five young, healthy volunteers, recruited amongst the students of the University of Applied Sciences Schmalkalden, completed two nights in a real-car driving simulator lab. Before the experiment, volunteers were trained in simulator driving and were introduced to the Compensatory Tracking Task (CTT). Also, they were instructed a) to maintain a regular sleep-wake schedule and to avoid any daytime napping, which was verified by actometric recordings, and b) to refrain from caffeinated drinks. Volunteers arrived at the lab at 10 pm. After wire-up (EEG, ECG), the experiment started at 11:30 pm. There were eight experiment sessions, each lasting one hour. The last session finished at 8:30 am. Volunteers were given a supervised break between 3:30 and 4:30 am, in which sleep was not permitted. Each session included a 40-minute driving session, a 10-minute CTT, and a 5-minute Psychomotor Vigilance Task (PVT). The PVT results are not discussed here. During the sessions, subjective ratings on the Karolinska Sleepiness Scale (KSS) was utilized as fatigue or sleepiness measure. As an objective performance measure, the variation of lane deviation (VLD), and the variation of the tracking error (VTE) were calculated. ECG was recorded continuously. To allow sufficient recovery, both trial nights for each subject were at least seven days apart.

3. METHODS

The heartbeat-to-heartbeat (RR) time series (in milliseconds) were derived from the ECG data. The cardiovascular measures described below were calculated using a five minute moving window with time steps of one minute. To enable a comparison of the HR measures between the 40 minute driving time and the 10 minute CTT, the values of HR measures obtained from the moving windows were averaged. This resulted in a set of HR measures of the driving and tracking tasks for each session. The HR measures represent the time, frequency and state space domain. The HR measures in the time domain are: (1) Heart Rate in beats per minute (B/m); (2) HRV as standard deviation of the RR values in milliseconds (ms); (3) pNN50 as percentage of consecutive RR intervals more than 50 ms; (4) Zcross as the number of zero crossings in the given time window after z-transforming the RR series.

In the frequency-domain, the power spectral density (PSD) of the RR time series using the fast Fourier transform (FFT) was calculated. Based on the PSD result, the power values for different frequency bands were calculated. The commonly used frequency bands are very low frequency (VLF, 0-0.04 Hz), low frequency (LF, 0.04-0.15 Hz), and high frequency (HF, 0.15-0.4 Hz). All three frequency bands were normalized by the total power. A simple method to capture this nonlinear characteristic is the so-called Poincaré plot in state space domain. A Poincaré plot of RR intervals is composed of the two-dimensional graphical presentation of the correlation between consecutive RR intervals. An ellipse is fitted onto the so called line-of-identity at 45 degrees to the normal axis. The standard deviation of the data perpendicular to the line-of-identity is defined as SD1 in milliseconds. SD1 describes the short-term variability which is mainly caused by respiratory activity. The standard deviation along the line-of-identity denoted by SD2 in milliseconds describes long-term variability.

In addition to the ECG data, several measures which capture different facets of driver fatigue and performance were collected. KSS is a standard subjective, and independent, measure of sleepiness on a numeric scale between 1 and 10. VLD is an objective measure of driving performance. Following the driving sessions, the visual motoric coordination was tested by CTT lasting 10 minutes. The CTT is a two-dimensional tracking task involving an annulus as a target positioned in the center of a display, and a moving disk. The disk is distracted from the target by two forces, a Coulomb-like and a random force difficult to predict for the subject. The task is to keep the disk in the center of the target by moving a trackball. The position and velocity of the disk is recorded as a function of time. To obtain a performance measure for the CTT we defined the variation of tracking error (VTE) analog to driving performance measure VLD. In this manner two objective and comparable performance measures for driving and tracking can be obtained. The task of keeping a...
vehicle in the lane is largely a psychomotor task involving eye-hand coordination. The term ‘tracking-ability’ is sometimes applied to it, pointing to a close similarity between the driving and tracking task (CTT).

For each HR measure, the difference between the driving results and the CTT were calculated in order to obtain the discrimination ability of the measures for workload and fatigue. These individual differences are very important for the data analysis. Even if task demands for driving and CTT are equal for two subjects, their reaction to the demands and the difficulty of the task may very well differ. Therefore, all HR measures described above were first calculated on the subject level then normalized using a maximum-minimum scaling. All averaged measures based on the five subjects were obtained after the scaling procedure. This procedure was not applied to the HR difference measures. The difference measures for all subjects were averaged without previous scaling.

4. RESULTS

There was a general trend for all cardiovascular measures, whereby the differences between both tasks increases with increasing sleepiness and decreasing operator performance (fig. 3a and fig. 3b). Despite individual differences between the subjects, these results were valid for all five subjects. Furthermore, the overall results for some cardiovascula measures depended on the changing fatigue levels, while for others this was not the case.

4.1 Sleepiness and Performance Measures

Fig. 2 shows the average performance results of 5 subjects over two nights. All measures in fig. 2 were normalized to fit them into one graph. The subjective sleepiness (KSS) correlates very well with the error measure (VLD) for driving, and VTE for the CTT. The fatigue (KSS) and error levels (VLD, VTE) were low at the beginning of each night and high for the last three sessions of each night. Subjective sleepiness was higher and performance lower on average in the second night. This implies that the state of the driver is most impaired in the last three sessions of the second night, resulting in a performance decrease in both tasks (fig. 2). A comparison between the first four and the last four sessions is displayed in fig. 3a and fig. 3b. As expected, the KSS and the task errors expressed by VLD and VTE nearly doubled from the first to the last four sessions. There were no differences in the performance measures VLD and VTE.

4.2 Time Domain Measures

The time domain comparison of the cardiovascular measures between driving and CTT included the following parameters: HR, HRV, pNN50 and Zcross. Significant differences between driving and CTT can be detected only in the HR itself (p < 0.05). HR in beats per minute was significantly higher during the driving task than during the CTT. This was true despite the drastic changes in fatigue and performance of the subjects between the early and late sessions. It can be concluded that the driving task, even under very monotonous conditions, is more demanding than the CTT. Nevertheless, there is a slight decrease of the HR over the course of the night for both tasks. The correlations of HR during driving (CTT) with VLD (VTE) is -0.21 (-0.39). This trend was observed before and explained as a time on task effect. The decrease in performance and increase in sleepiness affects the other time domain measures (HRV, pNN50, Zcross) in different ways. In all cases the value during the CTT higher. This difference increased during the late sessions (Fig. 3b). For example, HRV showed a significant correlation with VLD (0.48) and with VTE (0.70). The measure ‘pNN50’ did not correlate at all with VLD, but did correlate with VTE. The difference (A pNN50) between driving and CTT showed a significant correlation (-0.55) with KSS. The measure ‘Zcross’ on the other hand showed a strong correlation (-0.86) with VLD, but no correlation with VTE.

4.3 Frequency Domain Measures

The most interesting measures were in the frequency domain. The measure VLF showed very strong correlations with VLD (0.89) and VTE (0.79). VLF was not suitable for the workload discrimination, but showed the influence of the circadian time of day on the fatigue and performance data. This circadian influence was equal for driving and the CTT, which is reflected by the VLF measure. Therefore, there is no difference in the VLF band. The LF band only showed a similarly strong correlation for both tasks with the opposite direction. LF correlated with VLD (0.69) and with VTE (-0.71). There was an increasing difference between driving and CTT for the late sessions. In addition, the difference ‘A LF’ correlated with KSS (0.63), pointing to a fatigue influence for this frequency band. The HF band showed only a
correlation for the driving performance VLD (-0.69). The measure HF reflected the influence of respiration on the heart, which played a more important role for the more demanding driving task.

4.4 State Space Measures
From the two measures in the state space, only the long-term variability SD2 showed correlation to VLD (0.55) and to VTE (0.73). The behavior of SD2 mirrored closely the measure HRV from the time domain. The short-term variability SD1 mirrored the behavior of the HR. The only difference to the HR parameter were the SD1 values of the CTT were larger than the SD1 values for the driving.

5. CONCLUSIONS
The data analysis strongly suggested that HR was by far the best measure to distinguish between driving and tracking tasks. The other measures described in the publication did not show the same, clear difference. The discrimination ability of the cardiovascular measures between driving and CTT are illustrated in fig 4 for the HR and HRV. The green and yellow bars in fig 4 show the average differences ΔHR=HR(drive)-HR(CTT) and ΔHRV=HRV(drive)-HRV(CTT), respectively. The upper row of numbers displays the real difference of HR in beats per minute (B/m) for each session averaged over all subjects. The lower row shows the difference of HRV in milliseconds (ms).

From the consistent difference in HR, it can be concluded that monotonous night driving is more demanding than the tracking task. This was the case for sessions of low fatigue as well as high fatigue. The overall result for ΔHR strongly depended on the changing fatigue levels. HRV on the other hand was clearly the better measure to indicate strong fatigue, but gave contradicting results regarding differences in the workload tasks. Negative ΔHRV values indicated a strong fatigue influence during the late sessions of the second night. This confirmed the similar findings in [5, 6]. Determining the mental workload as absolute value under the changing conditions of fatigue is in our opinion very difficult. The prediction of the relative workload differences between two similar tasks is still complex, but possible as demonstrated in [4]. Two monotonous workload tasks can be distinguished by basic analysis of multiple cardiovascular measures. In addition to the two time domain measures HR and HRV, the frequency domain measures are important. Especially, the VLF band allows understanding and separating of the circadian influence on the data. Despite individual differences between the subjects, these results were valid for all five subjects. Cardiovascular data, workload and fatigue can only be calculated and determined per individual. The individual differences of the heart, the driving and the tracking skills, as well as in the ability to endure fatigue, cannot be neglected. It is always a specific task performed by an individual (who is in a certain fatigue state) that leads to a specific performance result. Therefore, any data analysis has to take place first on the individual level. Then a good scaling procedure has to be applied to make the data between the individual subjects comparable. In this case the maximum-minimum scaling was used. After successful scaling the data can be averaged. The inclusion of the PVT results is intended as well. It is important to recognize that the wide range of cardiovascular measures are only meaningful for a relative comparison of different tasks, and thus can only be understood under the combined investigation of cognitive workload and fatigue.

The use of additional technologies during driving especially adds to the mental workload side. Assistance technologies can either support or distract the driver in or from the primary driving task. In any case, an assessment of mental workload under different fatigue conditions with objective measures should be part of the design of any in-car user interfaces.

Beyond the workload determination during driving, it would be important to identify the most dangerous situation of high workload and high fatigue in order to discriminate this condition from the other combinations of workload and fatigue. This would help to avoid critical incidences in air traffic control and other occupations where extremely high workload and fatigue occur at the same time.

6. REFERENCES