JLR Heart: Employing Wearable Technology in Non-Intrusive Driver State Monitoring. Preliminary Study*

Vadim Melnicuk, Stewart Birrell, Panos Konstantopoulos, Elizabeth Crundall and Paul Jennings

Abstract— **This paper presents the results from a preliminary study where a wearable consumer electronic device was used to assess driver's state by capturing human physiological response in non-intrusive manner. Majority of state of the art studies have employed medical equipment drivers' state evaluation. Despite the potential gain in road safety this method of measuring physiology is unlikely to be accepted by private vehicle consumers due to its invasiveness, complexity, and high cost. This study was aiming to investigate possibility of employing a consumer grade wearable device to measure physiological parameters related to cognitive workload in real-time while driving i.e., drivers' heart rate. Furthermore, validity of captured heart activity metrics was analyzed to determine if wearable devices could be embedded into driving at its current technological state. The driving context was reproduced in desktop driving simulator, with 14 participants agreeing to take part in the study (μ = 28, σ = 8.5 years). Drivers were exposed to various road types, including pure Motorway, Rural, and Urban scenario modes. An accident was simulated in order to generate sudden cognitive arousal and capture participants' physiological response to the generated distress. It was found that a smartwatch is capable of reliable heart activity tracking in driving context. The results, supporting the relationship between cognitive workload level, generated by various complexity driving tasks, and Heart Rate Variability, were also presented.**

I. Introduction

Despite recent advances in driver assistance and in-car safety systems the risk of serious accidents on the roads is still present; it was forecasted that the total number of road traffic deaths and injuries worldwide is expected to rise by 65 percent between 2000 and 2020, putting road accidents into the third leading cause of the global deaths and injuries by 2020 [1]. According to World Health Organization, current road safety efforts fail to match the severity of this problem [2]. Therefore, it is important to scientifically approach the road safety issues and propose an innovative solution that could help to reduce road accident possibility.

*Research supported by Engineering and Physical Sciences Research Council (EPSRC) and Jaguar Land Rover

V. Melnicuk is with WMG, University of Warwick, Coventry, CV4 7AL, UK (phone: +44 24 765 28028; e-mail: V.Melnicuk@warwick.ac.uk).

S. Birrell is an Assistant Professor at WMG, University of Warwick, Coventry, CV4 7AL, UK (e-mail: S.Birrell@warwick.ac.uk).

P. Konstantopoulos is a Visual Performance Technical Specialist at Jaguar Land Rover, Banbury Road, Gaydon, CV35 0RR, UK (e-mail: pkonsta1@jaguarlandrover.com).

E. Crundall is a Human Factors Specialist at Jaguar Land Rover, Banbury Road, Gaydon, CV35 0RR, UK (e-mail: ecrundal@jaguarlandrover.com).

P. Jennings is a Professor at WMG, University of Warwick, Coventry, CV4 7AL, UK (e-mail: Paul.Jennings@warwick.ac.uk).

There are many road accident-causing factors, with human error being significantly dominant. According to survey conducted by National Highway Traffic Safety Administration (NHTSA), human error is a cause of approximately 94 percent of observed road accidents [3]. The most detrimental human error types are: recognition error (i.e., inattention, internal and external distractions, 41%), decision error (i.e., driving aggressively, driving too fast, 33%), performance error (i.e., overcompensation, poor directional control, 11%), and non-performance error (i.e., fatigue or sleepiness, 7%) [4]. Many studies have presented techniques to acknowledge and reduce different types of human error by detecting and counteracting error-causing factors in real time. For instance, driver workload, especially visual and manual origin of it, was found to cause driving performance impairments and distraction from the primary task of driving [5]. This is mostly due to presence of complex secondary tasks such as, use of In-Vehicle Information Systems (IVIS) or Consumer Electronic Devices (CEDs) while driving (e.g., cellular phone conversations, route navigation control and etc.) [5], [6]. Likewise, cognitive workload level can be affected due to presence of complex secondary tasks. Since cognitive state of a driver cannot be directly observed, the metrics of cognitive workload have to be gathered through human physiological response, performance metrics or by means of subjective measures [7]. However, subjective measures are fundamentally not suitable for commercial use, because they are generally measured before or after a driving activity by means of self-reporting therefore, cannot provide driver state reference for real-time systems.

Physiological parameters were also found to correlate to the cognitive workload in previous studies [7]–[9]. The method has been used priviously to assess the workload associated with voice-based in-vehicle interactions (e.g., cellular conversations or voice comands), where performance metrics, such as glancing patterns, are not directly impacted [10]. The heart activity records were identified as being sensitive to cognitive workload. Mehler et al. [11] used mean Heart Rate (HR) to consistently estimate relative differences in secondary task periods from single task driving in both simulated and real world driving. Similarly, Heart Rate Variability (HRV), which is defined as "*variability of time durations between every successful heartbeat*", was found to be influenced by cognitive workload [9], [12]. Some other human physiological parameters, such as brain activity, perspiration, and body temperature, were found to be sensitive to cognitive workload. However, Paxion et al. [7] concludes that the heart activity is the most sensitive physiological measure to track cognitive workload. Kramer [8] has outlined some advantages of employing physilogical metrics in cognitive workload estimation, such as unobtrusive measurements (e.g., direct tracking through minituarised sensors), physiology can also provide information regarding drivers state during absence of performance activity, and finaly, mixure of physiological parameters can provide multidimentional measure of cognitive workload. Author also listed some of the disadvantages, such as necessity of specialised equipment, abdundance of noise, and influence by factors other than workload e.g., emotional state and enviromental impact.

The sensory technology, harware and software have significantly advanced since early 90's. By 2006 the non-intrusive physiological tracking became available through wearable CEDs (e.g., fitness trackers, smartwatches, and chest straps). Some of these devices are embedded with multiple sensors capable of measuring number of steps, distance walked, Global Positioning System (GPS) coordinates, HR, HRV, Galvanic Skin Response (GSR), and Peripheral Body Temperature (PST) in real-time. These devices are mostly used to track physiology in health and fitness context and there is not much reported use in driving context. The metrics of numerous physiological parameters captured by such devices could be further fused with performance and contextual measurements to provide more reliable and valid driver state evaluation in real-time. Furthermore, these devices could be employed in development of driver state cautious Human-Machine Interfaces (HMIs) or adaptive Advanced Driver Assistance Systems (ADAS). Hence, contribute to improvement of road safety. Therefore, this paper primarily aims to explore the reliability and validity of wearable CEDs to measure human physiology while driving. As well as investigate validity of driver's HR captured by a wrist-worn consumer electronic device and possibility of employing HRV in driver's cognitive workload level estimation. It was hypothesized that, (i) HR activity measured from a smartwatch will not be significantly different to a baseline reference, (ii) variability of task complexity in different road types (i.e., motorway, rural, urban etc.) will influence HR, and (iii) HRV captured by a wearable device can be used to estimate level of cognitive workload in driving context.

II. Background

This section includes the brief review of the literature that presents the potential for heart activity tracking in driving context for workload level estimation.

The mean HR was found to successfully indicate increased levels of workload [9], [11], [13]. HR is usually presented in a form of mean number of heart Beats Per Minute (BPM). HR can be measured by means of Electrocardiography (ECG) or Photoplethysmography (PPG). ECG measures electrical heart activity and detects each QRS complex that forms Inter-Beat Intervals (IBIs), or normal-to-normal (NN) HR intervals. Whereas, PPG derives HR by detecting fluctuations in blood volume through a light emitting sensor.

HRV describes the variations between consecutive heartbeats [14]. It was found to correlate with age, mental and physical stress and workload, as well as attention [15]. According to Lee & Park [16], increase in cognitive workload had no effect on HR, while decrease in HRV could be observed. The time-domain measures, such as square root of the mean of the sum of the squares of differences between adjacent NN intervals (RMSSD) and NNs over 50 ms count divided by the total number of all NN intervals (p50NN) can be used to investigate fluctuations of HRV.

Frequency-domain analysis of HRV allows splitting the parameter into three frequency ranges: 0.0033-0.04 Hz Very Low Frequency (VLF), 0.04-0.15 Hz Low Frequency (LF), and 0.15-0.4 Hz High Frequency (HF). LF, HF and a ratio between those frequencies are mostly used to assess frequency-domain response of HRV. Lomb-Scargle Periodogram is a preferred method for calculating HRV frequency spectrum because it can directly use unevenly sampled IBIs and is robust to missed heart beats [17]. In summary, LF was found to reflect mainly sympathetic PNS tone. In the other hand, the HF was found to reflect parasympathetic PNS and is mostly influenced by breathing or Respiratory Sinus Arrhythmia (RSA) [18].

III. Methodology

A. Device Selection

Intel's BASIS Peak was chosen for this experiment. It is capable of tracking HR, forming minute-long estimates of Heart Beats Per Minute (BPM). It is also capable of measuring GSR, PST. This is a greater selection of physiological parameters than any other wrist-worn CED could offer. It also provides data collection and analysis benefits. However, for the purpose of this paper only HR metrics were taken into consideration. In order to evaluate the validity of HR captured by BASIS Peak it was compared to values collected by well-established device – POLAR H7 Heart Monitor. It measures electrical heart activity and is capable of detecting time between every successful heartbeat IBI (often referred to as "RR" interval). It also conforms to European Council Directive 93/42/EEC of 14 June 1993 concerning medical devices.

B. Desktop Simulator User Trials

To evaluate how effective BASIS Peak smartwatch was at accurately tracking heart activity in a driving context a desktop driving simulator study was adopted. Reimer & Mehler [10] have demonstrated that physiological measures of workload can be productively modelled under simulated driving conditions. WMG's 3xD Desktop Driving Simulator for Intelligent Vehicles (Figure 1), consisting of Logitech G27 steering wheel, pedal set, gear lever, and three 22" screens, as well as, software developed by XPI Simulation Ltd., was used. The desktop driving simulation was chosen to provide safe and comfortable environment for participants taking part in the experiment. Driving scenario consisted of simulated accident on the road in order to capture hypothesized human physiological response to a highly distressful and attention-demanding situation; therefore, real world driving was not suitable for such driving scenario setup. Likewise, fully controlled and repeatable driving scenario could not be implemented in real world, because there is a risk of uncontrolled events appearing during an experiment, which might influence driver's state. The simulator allowed us to produce fully controlled driving scenario, where every surrounding vehicle was scripted to perform predefined movement sequences without any variance in environmental setup or traffic level.

Figure 1: Desktop driving simulator.

Participants were required to wear BASIS Peak smartwatch on their wrist of their preferred hand. A sweatband was added on top of the device to secure the position, cover it from any direct sunlight, as well as, reduce participants' visual distraction caused by the device's graphical display. Next, POLAR H7 Heart Monitor was fitted around a participant's chest. The Heart Monitor requires direct access to the skin and its electrodes to be slightly moisturized. Once both devices were fitted, the participants were shown to the desktop driving simulator and guided through the setup. After that, familiarization with in-car interior and simulator controls took place. It was indicated that simulated vehicle has a dynamic model of a five-door Peugeot 206, which had an automatic transmission. The closed-loop practice scenario was loaded to allow participants to familiarize with vehicle controls. After a participant felt confident to continue the experiment, the experimental scenario was loaded. The capturing of resting BPM was initiated for the period of five minutes. Participants were asked to sit calmly for five minutes. Although, readings of resting HR were taken not in a supine position, it was ensured that participants are free from any distraction and manual activity.

C. Participants

The recruitment of participants was conducted internally at University of Warwick. The call for participants was disseminated using internal email advertisement and notice board posters put throughout the department. Participants were required to contribute 60 minutes of their time, and meet the following criteria: be over 21 years of age; hold full category "B" driving license; have normal or correct-to-normal vision; and do not have any cardiovascular diseases. Suitable volunteers were issued with a Participation Information Leaflet, which provided an in-deep explanation of experiment methodology, risks involved, as well as, advantages and disadvantages of participation. Trials lasted for two and a half weeks from 13th to 29th of April 2015 with 14 participants agreeing to take part. The sample consisted of participants of mixed age ($\mu = 28$, $\sigma = 8.5$ years).

D. Scenario Design

The driving scenario was designed to allow participants to experience different levels of workload. This was achieved by incorporating three discrete types of roads, traffic levels, primary driving tasks, frequency of auditory navigation commands, and an accident simulation (see TABLE I). The assumption was made that variation in road complexity and demand have increased levels of visual, auditory, manual, and presumably cognitive workload.

TABLE I: Scenario Design Summary.

Scenario modes	Duration (min)	Road layout	Traffic Density	Speed Limit (mph)
Resting	5	N/A	N/A	N/A
Transition	7	Variable	Variable	Variable
Urban	5	One lane with in-front vehicle present	Heavy traffic	30
Rural	7	One empty lane	Absence of traffic	60
Motorway	10	Three empty lanes	Absence of traffic	70
Accident	1	One empty lane	One in-front vehicle	60

The minute when an interchange between road types (scenario modes) occurred was labelled as "Transition". During the "Accident" an in-front vehicle was scripted to emergency stop rapidly decelerating to zero, causing a participant to react to the event i.e., avoid a crash by steering away or crash into the vehicle. This has, assumingly, caused participants to experience high arousal and sudden increase of workload level. Voice navigation commands were embedded into the scenario to ensure that subjects follow the pre-planned route, and were put before every junction and roundabout. The study researcher was present in the simulation room during the experiment to make sure that a participant follows the route correctly and answer any study related questions. Overall, the experimental driving scenario lasted for approximately 35 minutes, depending on driver's performance.

E. Data Collection

The quantitative data sets (BPM, GSR, and PST) were collected throughout the desktop simulator user trials. These consist of captured sensory data from two wearable devices, BASIS Peak and POLAR H7. Data collection was initiated through two separate mobile devices. LG Nexus 7 tablet with preinstalled "HRV Expert by CardioMood" mobile application was synchronized with POLAR H7 and iPhone 6 with preinstalled "Basis Peak" application was synchronized with BASIS Peak smartwatch. The quantitative data, participants' answers to pre-experiment questionnaire, was collected at the beginning of every experiment in order to determine participants' demographics, such as gender and age. Due to nature of this study i.e., benchmarking of consumer grade devices for physiological tracking in driving context, no additional participants' information was recorded.

F. Data preparation

The raw heart activity data recorded through POLAR H7 had to be averaged to form minute long measures of BPM. This was implemented in order to mirror the data format of BASIS Peak, which records average BPM per every minute. Additionally, BASIS Peak does round BPM values to the nearest whole number. We are not aware how this rounding is performed thus, this step was not added to the averaging algorithm for POLAR. The individual data sets recorded by POLAR H7 were run through BPM averaging script prepared in MATLAB. Next, data was imported into Excel spreadsheet for measurement times from both devices to be synchronized. Finally, the data set was imported into IBM SPSS Statistics 22 for an analysis.

The HRV parameters were derived from RR intervals recorded by POLAR H7 during the user trials. Following the methods of deriving time-domain and frequency-domain HRV parameters, discussed in "Background" section, MATLAB script was written. The script continuously calculated time-domain measures: RMSSD of the last 10 data points, as well as, RMSSD and p50NN of the last 100 data points. It also continuously calculated HRV frequency spectrum using Lomb-Scargle method. The spectrum was split into the following frequency ranges: VLF, LF, HF, and LF to HF ratio of the last 100 data points. All derived parameters were exported into a separate comma-separated values (.csv) file.

IV. Results

A. Participants

Two participant's data sets were corrupted and were excluded from the analysis. This was detected during initial visual evaluation of the acquired data. The continuous HR line graphs were plotted for every individual. One graph showed a significant data loss. This could happen due to displacement of POLAR H7 Heart Monitor during the trial. Another graph revealed lack of fluctuations of HR with unusual presence of artefacts (measures were out of maximum HR range). Therefore, initial sample containing 481 data sets was reduced to 406 (12 participants' data).

B. Benchmarking BASIS against POLAR

Mean BPM for the BASIS Peak, calculated over the entire duration of the driving scenario including all participants' data, was 68.94 (σ = 10.35) and for the POLAR H7 it was 69.99 (σ = 10.90). The absolute mean BPM for POLAR H7 (non-averaged raw BPM values), which was 71.66 (σ = 11.87) was also derived. It differs to an averaged value by 1.67. This demonstrates that averaging an average can sometimes lead to variation for this type of analysis. Despite presence of an averaging mistake, mean BPM of per-minute-average was used in further analysis to allow assessment of measurement differences between two devices.

Neither measurement method (BASIS or POLAR) reveals unusual features, such outliers (Figure 2). The mean BPM differences between two measurements methods (POLAR minus BASIS) were not normally distributed (Shapiro-Wilk test, $p<0.001$), with distribution negatively (-0.3) skewed (Figure 3). Despite the mean difference between per-minute averaged readings being approximately one BPM, it was highly significant (Wilcoxon signed-rank test, Z = -10.47, $p<0.001$). This may be present due to BASIS Peak's added BPM rounding step. However, there was a strong correlation between BPM readings captured by BASIS Peak and POLAR H7 (Pearson's Correlation = 0.978, $p<0.01$).

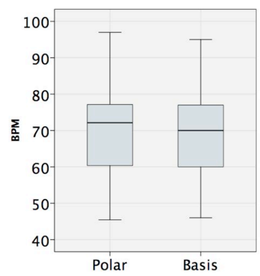

Figure 2: Heart Rate summary box plot for both measurement methods.

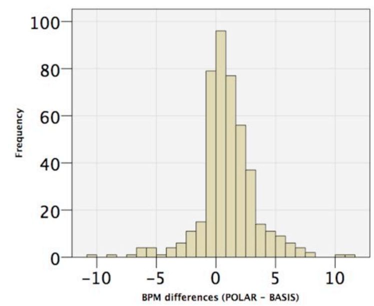

Figure 3: Histogram of BPM differences between two measurement methods (POLAR minus BASIS).

The impact of road types on BPM readings captured by BASIS and POLAR (averaged per-minute measurements) was investigated using univariate analysis of variance (ANOVA). Results suggest that BPM variance in both measurement methods was not significantly affected (BASIS: F=0.578, p=0.717; POLAR: F=0.476, p=0.794) by task complexity variations in different road types. However, raw BPM metrics, collected using POLAR H7, showed significant difference in various road types (ANOVA, F=31.339, $p<0.001$). The measurement methods were further visually compared using bar plot of BPM means categorized by driving scenario modes (Figure 4). The mean BPM tends to lower (although not significantly) with decreasing complexity of driving scenario in both measurement methods.

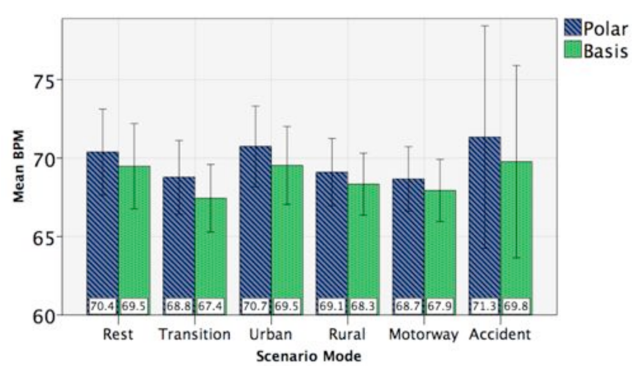

Figure 4: Mean BPM measurements categorized by scenario mode.

Furthermore, Scatter Plot of POLAR against BASIS measurement categorized by road type (Figure 5) was plotted to assess regression of metrics. Linear regression goodness-of-fit measures ($R_{2\ URBAN}$ = 0.933, $R_{2\ RURAL}$ = 0.961, $R_{2\ MOTORWAY}$ = 0.981) are also affected by increasing complexity of driving between road types. Regression tends to lower with increased level of workload ($R_{2\ URBAN} < R_{2\ RURAL} < R_{2\ MOTORWAY}$).

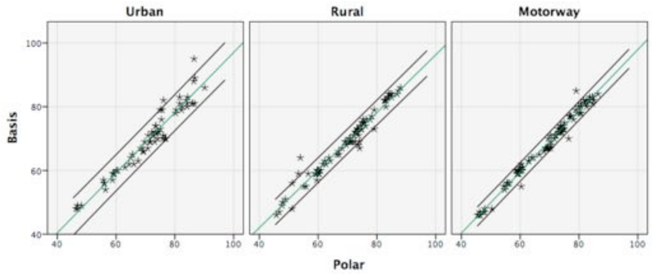

Figure 5: POLAR against BASIS measurements categorized by road type.

C. Heart Rate Variability

Figures 6 and 7 represent the mean affect of scenario mode complexity on longer-term HRV. In total 13980 RMSSD and p50NN of 100 IBI data points, continuously calculated over "Rest", "Motorway", "Rural" and "Urban" scenario modes, were analyzed. The impact of road types on HRV readings was investigated using ANOVA. Results suggest that RMSSD and p50NN variance was significantly affected (RMSSD100: F=96.61, p<0.001; p50NN: F=34.50, p<0.001) by variations of complexity in different scenario modes. The metrics were further analyzed using post-hoc pairwise comparison with Bonferroni correction. The variance in RMSSD100 and p50NN100 is not significantly different only between "Motorway" and "Rural" road types (RMSSD100: p=0.900; p50NN100: p=0.600). The rest of the road types demonstrated significant difference in variance between each other (p<0.001).

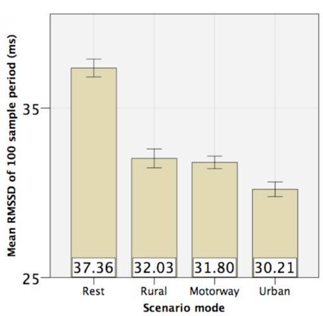

Figure 6: Mean RMSSD of 100 sample period categorized by scenario mode.

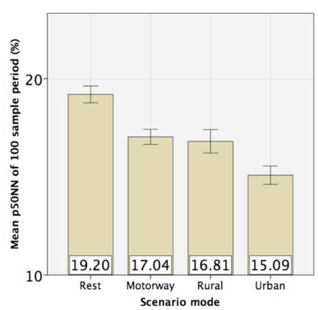

Figure 7: Mean p50NN of 100 samples period categorized by scenario mode.

Figure 8 represents the mean affect of scenario mode complexity on short-term HRV. In total 19040 RMSSD of 10 IBI data points, continuously calculated over "Rest", "Motorway", "Rural", "Urban", and "Accident" scenario modes, were analyzed. The impact of road types on short-term HRV was investigated using ANOVA. Results suggest that RMSSD10 variance was significantly (F=44.23, p<0.001) affected by variations of complexity in different scenario modes. The RMSSD10 was further analyzed using post-hoc pairwise comparison with Bonferroni correction. The variance was shown not significantly different in "Motorway-Rural" (p=1.000) and "Urban-Accident" (p=1.000) cases. The rest of the scenario modes demonstrated significant difference in variance between each other.

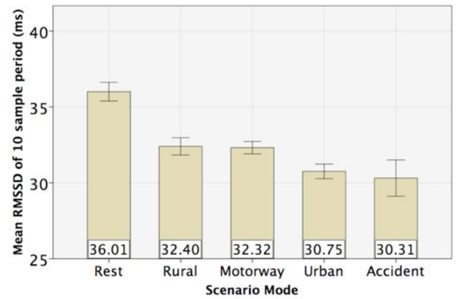

Figure 8: Mean RMSSD of 10 sample period categorized by scenario mode.

HRV was further studied using its frequency-domain measures. MANOVA was used to study an impact of road types on HRV frequency-domain measures. Results suggest that VLF, LF, and HF variance was significantly affected (Pillai's trace: F=255.45, p<0.001) by variations of complexity in different scenario modes. The post-hoc pairwise comparison with Bonferroni correction was applied to study differences in variance between scenario modes. The insignificance in variance was found between "Motorway" and "Rural" scenario modes in all frequency ranges (VLF: p=1.000; LF: p=0.152; HF: p=1.000). Additionally, the variance between "Rural" and "Rest" scenario modes in VLF range was also found insignificantly different (p=0.056). The LF to HF ratio was significantly affected by varying driving complexity in different scenario modes (F=263.619, p<0.001) (Figure 9).

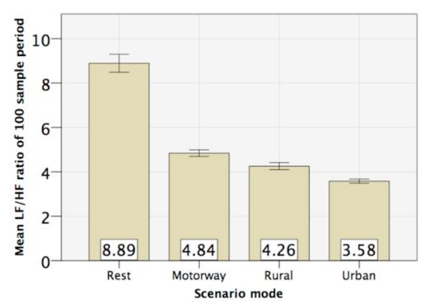

Figure 9: Mean LF/HF ratio of 100 sample period categorized by scenario mode.

V. Discussion and Conclusions

This preliminary study has addressed the need for innovative, non-intrusive and cost effective driver state monitoring solution that could help to estimate levels of workload in driving context. The feasibility of workload estimation through driver's physiology has been confirmed previously in the literature however, most of the studies have employed medical equipment to capture human's physiology for workload level estimation [15], [16]. This report includes results produced to support a potential use of a smartwatch as a tool to capture driver's physiology in real-time. CEDs are mostly used to track physiology in health and fitness context and there is not much reported use in driving context.

Our first hypothesis is that HR measured from a smartwatch will not be significantly different to a baseline reference. The validity of HR measurements was compared to the ones measured by POLAR H7 Heart Monitor, a well-established method of capturing electrical heart activity. According to the results, BASIS Peak smartwatch was found to reliably track HR. Although, two measurement methods did not agree, HR from BASIS was consistently just one BPM out across scenario modes and it demonstrated significant correlation with metrics measured by POLAR H7. No bias towards any particular BPM range or scenario mode was found. Therefore, we can accept Hypothesis "1" and say with some confidence that BASIS Peak could be employed in relatively accurate HR tracking in driving context.

The affect of task complexity variation in different scenario modes on mean HR was also investigated. Results suggest that continuous minute-long BPM variance in both measurement methods was not significantly affected by task complexity variations in different road types. In contrast, the raw BPM values measured by POLAR H7 were found significantly different in various scenario modes. Perhaps, the BPM averaging stage has influenced the data by smoothing out some of the essential HR fluctuations. Besides, the data-averaging step significantly reduces number of data points and leads to wide confidence limits. Thus, Hypothesis "2" can be accepted according to analysis of raw BPM metrics from POLAR H7, but must be rejected in the case when per-minute averaging of BPM is performed (i.e., BASIS Peak's averaging).

In order to test Hypothesis "3", the HRV was analyzed using time-domain and frequency-domain methods. The mean RMSSD and p50NN were studied across all scenario modes to identify any dependencies on task complexity. Both measures were found to be significantly different across most of the scenario modes. The short-term (10 IBIs) and long-term (100 IBIs) HRV was found to be very similar and insignificantly variant in "Motorway" and "Rural" scenario modes. This is an expected behavior for HRV and perhaps, occurred due to design similarities in those scenario modes. The HRV was also found to follow driving task complexity pattern. According to the literature time-domain measures of HRV tend to decrease with increased levels of workload [9]; this trend was confirmed in our findings. Similar trend was observed in frequency-domain metrics as well as similarities in "Motorway" and "Rural" scenario modes. Therefore, Hypothesis "3", which states that HRV can be used to estimate level of cognitive workload in driving context, can be accepted for both time-domain and frequency-domain measures.

In summary, this paper presents supporting evidence for adopting wrist-worn consumer grade device for the purpose of driver state monitoring. However, some limitations persist. It is necessary to capture all RR intervals in order to perform HRV analysis. None of the commercially available smartwatches are able to provide such data rate yet. In contrast, Renevey et al. [19] have demonstrated a proprietary wrist-worn devices capable of measuring RR intervals in real-time. The future studies aiming to embedded wrist-worn CEDs for physiological tracking in driving context should address this limitation. Future research should also focus on the following areas: (i) investigate user acceptance of physiological tracking by means of CEDs for safety purposes in driving context, (ii) investigate a possibility of embedding CEDs in real world driving from technological and human factors point of view, and (iii) further investigate potential for workload level estimation using HRV.

Acknowledgment

The support for this work was provided by Engineering & Physical Sciences Research Council (EPSRC) and Jaguar Land Rover (JLR).

References

[1] E. Kopits and M. Cropper, "Traffic fatalities and economic growth," *Accid. Anal. Prev.*, vol. 37, no. 1, pp. 169–178, 2005.

[2] World Health Organisation, "World report on road traffic injury prevention," 2004.

[3] NHTSA, "Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey," Washington, US, 2015.

[4] NHTSA, "National Motor Vehicle Crash Causation Survey Report to Congress," 2008.

[5] S. a. Birrell, M. Fowkes, and P. a. Jennings, "Effect of using an in-vehicle smart driving aid on real-world driver performance," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 4, pp. 1801–1810, 2014.

[6] T. C. Lansdown, N. Brook-Carter, and T. Kersloot, "Distraction from multiple in-vehicle secondary tasks: vehicle performance and mental workload implications.," *Ergonomics*, vol. 47, pp. 91–104, 2004.

[7] J. Paxion, E. Galy, and C. Berthelon, "Mental workload and driving," *Front. Psychol.*, vol. 5, no. December, pp. 1–11, 2014.

[8] A. F. Kramer, "Physiological metrics of mental workload: A review of recent progress," 1990.

[9] D. De Waard, "The Measurement of Drivers ' Mental Workload," University of Groningen, 1996.

[10] B. Reimer and B. Mehler, "The impact of cognitive workload on physiological arousal in young adult drivers: a field study and simulation validation," *Ergonomics*, vol. 54, pp. 932–942, 2011.

[11] B. Mehler, B. Reimer, and Y. Wang, "Comparison of heart rate and heart rate variability indices in distinguishing single task driving and driving under secondary cognitive workload," *Proc. Sixth Int. Driv. Symp. Hum. Factors Driv. Assessment, Training, Veh. Des.*, pp. 590–597, 2011.

[12] A. Hoover, A. Singh, S. Fishel-Brown, and E. Muth, "Real-time detection of workload changes using heart rate variability," *Biomed. Signal Process. Control*, vol. 7, no. 4, pp. 333–341, 2012.

[13] K. a. Brookhuis and D. de Waard, "Monitoring drivers' mental workload in driving simulators using physiological measures," *Accid. Anal. Prev.*, vol. 42, no. 3, pp. 898–903, 2010.

[14] M. Malik, "Heart rate variability. Standards of measurement, physiological interpretation, and clinical use," *Eur. Heart J.*, vol. 17, pp. 354–381, 1996.

[15] G. Berntson, T. Bigger, D. Eckberg, P. Grossman, P. Kaufmann, M. Malik, H. Nagaraja, S. Porges, P. Saul, P. Stone, and M. van der Molen, "Heart rate variability: Origins, methods, and interpretive caveats," *Psychophysiology*, vol. 34, no. 6, pp. 623–648, Nov. 1997.

[16] D. H. Lee and K. S. Park, "Multivariate analysis of mental and physical load components in sinus arrhythmia scores.," *Ergonomics*, vol. 33, no. 1, pp. 35–47, 1990.

[17] J. a. Healey and R. W. Picard, "Detecting Stress During Real-World Driving Tasks Using Physiological Sensors," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 2, pp. 156–166, Jun. 2005.

[18] L. K. McCorry, "Physiology of the autonomic nervous system," *Am. J. Pharm. Educ.*, vol. 71, no. 4, 2007.

[19] P. Renevey, J. Solà, P. Theurillat, M. Bertschi, J. Krauss, D. Andries, and C. Sartori, "Validation of a wrist monitor for accurate estimation of RR intervals during sleep," in *35th Annual International Conference of the IEEE EMBS*, 2013, pp. 5493–5496.