

Developing an Objective Indicator of Fatigue: An Alternative Mobile Version of the Psychomotor Vigilance Task (m-PVT)

Michael Scott Evans¹[0000-0003-0901-2190], Daniel Harborne²[0000-0002-2376-8341] and Andrew P. Smith¹[0000-0001-8805-8028]

¹ School of Psychology, Cardiff University,
63 Park Place, Cardiff, CF10 3AS, Wales, United Kingdom
EvansMS3@cardiff.ac.uk

² School of Computer Science and Informatics, Cardiff University,
Cardiff, Wales, UK

Abstract. Approximately 20% of the working population report symptoms of feeling fatigued at work. The aim of the study was to investigate whether an alternative mobile version of the ‘gold standard’ Psychomotor Vigilance Task (PVT) could be used to provide an objective indicator of fatigue in staff working in applied safety critical settings such as train driving, hospital staffs, emergency services, law enforcements, etc., using different mobile devices. 26 participants mean age 20 years completed a 25-minute reaction time study using an alternative mobile version of the Psychomotor Vigilance Task (m-PVT) that was implemented on either an Apple iPhone 6s Plus or a Samsung Galaxy Tab 4. Participants attended two sessions: a morning and an afternoon session held on two consecutive days counterbalanced. It was found that the iPhone 6s Plus generated both mean speed responses (1/RTs) and mean reaction times (RTs) that were comparable to those observed in the literature while the Galaxy Tab 4 generated significantly lower 1/RTs and slower RTs than those found with the iPhone 6s Plus. Furthermore, it was also found that the iPhone 6s Plus was sensitive enough to detect lower mean speed of responses (1/RTs) and significantly slower mean reaction times (RTs) after 10-minutes on the m-PVT. In contrast, it was also found that the Galaxy Tab 4 generated mean number of lapses that were significant after 5-minutes on the m-PVT. These findings seem to indicate that the m-PVT could be used to provide an objective indicator of fatigue in staff working in applied safety critical settings such as train driving, hospital staffs, emergency services, law enforcements, etc.

Keywords: Psychomotor Vigilance Task (PVT), Mental Workload, Occupational Fatigue, Objective Indicator of Fatigue, Attention.

1 Introduction

In order to be able to meet task demands, there is usually a required amount of operator resources needed, referred to as human mental workload [1]. According to Hart

and Staveland [2], human mental workload can be defined as a ‘cost incurred by a human operator to achieve a particular level of performance’ and evolves from interactions between task demands, circumstances, skills, behaviour, and perceptions.’ Therefore, human mental workload – often referred to as cognitive load – can be intuitively defined as the amount of mental work necessary for a person to complete a task over a given period of time [3, 4]. However, nowadays human mental workload is more generally defined as the measurement of the amount of mental resources involved in a cognitive task [5].

Human mental workload can be measured in real time using a variety of psychological and physiological techniques, which include; subjective psychological self-reported measures e.g., the NASA Task Load Index (NASA-TLX) [2, 6–8] and the NASA-MATB (National Aeronautics and Space Administration Multi-Attribute Task Battery [9] as well as objective physiological measures e.g., heart rate (HR), galvanic skin response (GSR), body temperature, electrocardiogram (ECG), electroencephalogram (EEG), and eye tracking [8, 10–19], and which have been extensively examined in various safety critical environments including; aviation [7, 20], train driving [21], car driving [22–24], and in an operating theater [6] but to name a few.

According to Wickens [25], the greatest value of conducting scientific human mental workload research is to be able to predict the consequences of high mental workload on performance. In other words, to better understand an individual's decision to consciously engage in a safe behaviour or in a potentially dangerous behaviour that could have devastating consequences. As a result, the concept of human mental workload has long been recognised as an important factor in individual performance [26–29]. Xie and Salvendy [29] state that both underload (i.e., low mental workload) and overload (i.e., high mental workload) degrade performance, whereby high and low levels of human mental workload have been shown to lead to operator error [22]. Longo [3] outlines that during low mental workload, individuals are more likely to experience levels of frustration and annoyance when processing information, which could result in an increase in their reaction time (RT). In contrast, during high mental workload, individuals could experience confusion, which may result in a decrease in their information processing capacity, which could directly increase the likelihood of errors and mistakes. Therefore, these low and high mental workload information processing stages could have potentially dangerous consequences, especially in safety critical environments. Byrne [30] points out that the main application of mental workload has been to investigate situations where cognitive demand exceeds the acceptable safety tolerance threshold so that workload can be effectively reduced. Therefore, in high risk safety critical environments, the measurement of mental workload is of upmost importance due to its potential implications [31]. However, Xie and Salvendy [29] identified that the effect of fatigue on mental workload is not often considered in human mental workload research. Nevertheless, research carried out by Smith and Smith [32] on conductors/guards and engineers from the rail industry who work in high risk safety critical environments found that workload increased fatigue. However, subjective measures were predominately used in Smith and Smith's study. As a result, there is a need for an alternative mobile objective indicator of fatigue that can be used in high

risk safety critical environments. In a controlled laboratory setting, the human Psychomotor Vigilance Task (PVT) [see 33, 34, for review] has become the widely accepted ‘gold standard’ tool for assessing the impact of fatigue on human cognitive neurobehavioral performance for monitoring temporal dynamic changes in attention [35–38]. The aim of the study was to investigate whether an alternative mobile version of the ‘gold standard’ Psychomotor Vigilance Task (PVT) could be used to provide an objective indicator of fatigue in staff working in applied safety critical settings, such as train driving, hospital staffs, emergency services, law enforcements, etc.

The rest of the paper is organised as follows. Section 2 describes related work on the Psychomotor Vigilance Task (PVT) while also extracting relevant studies to identify the gaps and rationale for the need of an alternative objective indicator of fatigue in staff working in applied safety critical settings. Section 3 outlines the design and empirical methodology of the proposed alternative mobile Psychomotor Vigilance Task (m-PVT). Section 4 presents the empirical results and discussion of the m-PVT. Finally, Section 5 provides a critical conclusion of the proposed alternative m-PVT and suggestions for future work.

2 Related Work

The Psychomotor Vigilance Task (PVT) can be traced back from the early work in simple reaction time (SRT) studies that were carried out by Wilhelm Maximilian Wundt (1832 – 1920) and continued by James McKeen Cattell (1860 – 1944) [39]. It is important to note that the modern PVT has been refined several times over the years [40–42] from its original development by Dinges and Powell [33] and has been shown to be sensitive to sleep deprivation, fatigue, drug use, and age. The PVT has also been widely implemented using a handheld device known as the PVT-192 (Ambulatory Monitoring Inc., Ardsley, New York, USA), as well as being extensively validated by various researchers [40, 43–47].

According to Basner, Mcguire, Goel, Rao and Dinges [48] and Basner et al. [49], the PVT-192 records participants’ sustained attention based on repeated reaction time (RT) trials to visual stimuli that occur at random inter-stimulus intervals (ISI) that are between 2–10 seconds, for a standard 10-minute period. In summary, the PVT-192 device operated by presenting participants with a stimulus that consisted of a four-digit millisecond counter that appears in a light-emitting diode (LED) dot-matrix display. The response consisted of a left or right button press, which depended on the configuration of the PVT-192 setup. The time difference between the stimulus presentation and the response constituted the participant’s reaction time (RT). Each RT value was stored in the device and then uploaded to a personal computer, where the individual RTs are post-processed with the REACT software (Ambulatory Monitoring Inc., Ardsley, New York, USA), or other commercially available software, into summary statistics, such as the mean RT or the mean number of lapses (RTs ≥ 500 milliseconds) per session [33, 40, 48, 50, 51]. For example, in Roach, Dawson, and Lamond’s study [45], each participant performed either 5 minutes or 10 minutes RT sessions spaced at predetermined intervals (e.g., every 2 hours) for a prolonged duration (e.g., 28 hours), where each

session consists of either 50 trials (equivalent to 5 minutes), or 100 trials (equivalent to 10 minutes). However, Khitrov et al. [52] tested the average delay of the PVT-192 and found that the recorded delay was greater than what was stated by the PVT-192 manufacturer. The delay recorded by the researchers was on average 2.4 ms greater when compared to the manufacturer's reported delay of 1 ms. Nevertheless, it is important to highlight that Khitrov et al. [52] did acknowledge the possibility that the difference found could have been due to the non-instantaneous nature of the light detection circuit, or the actual delay associated with the PVT-192, since their experimental design did not permit them to be able to distinguish between these possibilities.

Dinges and Powell [33] have shown that the 10-min PVT is highly reliable. Roach, Dawson and Lamond [45] wanted to investigate whether 90 seconds could also be sufficiently sensitive enough to detect the effects of fatigue in comparison to their earlier research [see 43, for review], where they were able to find significant fatigue-related impairment during the first 5-min of a 10-min PVT. In this study, the researchers compared participants' neurobehavioral performance using the PVT between three different time durations (90 seconds, 5-min, and 10-min) to identify whether a shorter PVT could also be sensitive enough to detect the effects of fatigue. They found that it was only possible to implement a 5-min PVT as a substitute of the 10-min PVT, and not a 90 seconds PVT, thus only further supporting their earlier research [43]. However, it is important to note that analyses of their study were carried out using the mean RT and not the mean speed response (1/RT). Basner and Dinges [43] have identified that the mean RTs should not be the primary measure of alertness, and instead considering using the alternative primary measure of 1/RTs. In a later study, Basner, Mollicone and Dinges [42] aimed to further shorten the 5-min PVT [45] by developing a modified 3-min version of the PVT (PVT-B). They found that this 3-min version could be a useful tool for assessing behavioural alertness in settings where the 'gold standard' 10-min PVT could be more difficult or impractical to implement due to the nature of the study or location. However, further validation is required to determine whether both the 5-min PVT and PVT-B versions could indeed be sensitive enough to detect reduced levels of fatigue. Therefore, this study aimed to investigate a mobile version of the Psychomotor Vigilance Task (m-PVT) that could also be used to provide an objective indicator of fatigue in staff working in applied safety critical settings such as train driving, hospital staffs, emergency services, law enforcements, etc.

3 Design and Methodology

The aim of the study was to investigate whether an alternative mobile version of the 'gold standard' Psychomotor Vigilance Task (PVT) could be used to provide an objective indicator of fatigue in staff working in applied safety critical settings such as train driving, hospital staffs, emergency services, law enforcements, etc. The study received ethics approval from Cardiff University's Ethics Committee (EC.16.02.09.4464R). The study conformed to the seventh amendment of the Declaration of Helsinki 1964 [53] and all participants gave their informed written as well as electronic consent following the explanation of the nature of the study in written form.

3.1 Participants

26 (3 male and 23 female) participants with a mean age of 20 years ($SD = 1.66$) were recruited as volunteers from Cardiff University via the Experimental Management System (EMS) to take part in the study. The study involved participants attending two sessions, a morning session (i.e., before 11:00) and an afternoon session (i.e., after 17:00), which were held on two consecutive days and counterbalanced, in exchange for £10. The study lasted 60 minutes in total for both sessions.

3.2 Materials / Apparatus

The mobile Psychomotor Vigilance Task (m-PVT) was presented to participants on one of two mobile devices: Apple's iPhone 6s Plus running Apple's iOS version 9.3.1 (Apple Inc.) or Samsung's Galaxy Tab 4 (Samsung Electronics Co. Ltd.) running on Android's operating system (OS) version 4.4.2 KitKat (Alphabet Inc.). The m-PVT ran in the following hardware configurations for the iPhone 6s Plus: system chip (Apple A9 APL1022), processor (Dual-core, 1840 MHz, Twister, 64-bit), graphics processor (PowerVR GT7600), and system memory (2048 MB RAM), and for the Samsung Galaxy Tab 4: system chip (Marvell PXA1088), processor (Quad-core, 1200 MHz, ARM Cortex-A7), graphics processor (Vivante), and system memory (1536 MB RAM). The iPhone 6s Plus had the following hardware configurations: the m-PVT was displayed on either a 5.5-inch (diagonal) 1920×1080 -pixel native resolution at 401 ppi Retina high definition display (iPhone 6s Plus), or a 7-inch (diagonal) 1280×800 -pixel (WXGA) native resolution at 216 pixels per inch (ppi) liquid crystal display (LCD) display (Samsung Galaxy Tab 4).

The m-PVT was programmed using the client code HTML, and CSS for the page visualisation and layout. JavaScript was also used to initiate the m-PVT, which was run using the Dolphin Web Browser (MoboTap Inc.) on both an Apple's iPhone 6s Plus and Samsung Galaxy Tab 4 (Dolphin Web Browser versions; Apple app version 9.9.0, and Android app version 11.5.6, respectively). The rationale for selecting the Dolphin Web Browser for this study was that it allowed the full screen feature to be enabled across the two different operating systems (OS), Apple iOS and Android OS platforms for both mobile devices. Other more native mobile internet browsers of each OS platform, such as Safari (Apple) and Chrome (Android) including Firefox, to name a few, did not permit full screen. Qualtrics Surveys (Qualtrics Labs, Inc.) were also used to collect demographic information from participants. These surveys were also implemented on both Apple's iPhone 6s Plus (iOS app version 13.28.06) and Samsung Galaxy Tab 4 (Android app version 1.0.38).

3.3 Statistical Analyses

IBM's Statistical Package for the Social Sciences (SPSS) version 23 for Mac was used to analyse the data. A combination of various statistical procedures were carried out on the data; descriptive analyses, mixed-design analysis of variance (ANOVA) and a two-

way analysis of variance (ANOVA) to further explore interactions. The level of $\alpha < .05$ was used for all statistical tests of this experiment.

3.4 Design

The experiment employed a $2 \times 2 \times 6$ mixed-design analysis of variance (ANOVA) with mobile device (Apple's iPhone 6s Plus or Samsung's Galaxy Tab 4) as the between-subjects factor, \times time of day (Morning or Afternoon) \times time on task (1-minute; 5-minutes; 10-minutes; 15-minutes; 20 minutes; or 25-minutes) as the within-subjects factors. The morning session (i.e., before 11:00) and the afternoon session (i.e., after 17:00) were held on two consecutive days and counterbalanced.

3.5 Procedure

In order to ensure participants were fully aware of the inclusion and exclusion criteria, all participants were contacted using Cardiff University's Experimental Management System (EMS) emailing system 48 hours prior to participation and further reminded 24 hours before the start, in addition to being provided with brief instructions through EMS.

The study was administered using mobile devices. Participants were either assigned to using an iPhone 6s Plus or a Samsung Galaxy Tab 4. To increase validity and standardisation, all instructions were administered to participants in written form for both the morning and the afternoon session. This study consisted of two parts. The first part was the mobile Psychomotor Vigilance Task (m-PVT) reaction time test, which was a modified version of the Dinges and Powell's [33] Psychomotor Vigilance Task. The m-PVT was run on the Dolphin Web Browser mobile application. The second part was the demographic questionnaire that was distributed within Qualtrics Surveys mobile application. In this modified version, the mobile Psychomotor Vigilance Task (m-PVT) (see Figure 1), participants were presented with on-screen instructions and a button at the end that read 'Start'. In each trial, participants were shown a black screen background, and at the centre of the screen they would be presented with a large red fixation circle. The red fixation circle (i.e., inter-stimulus interval) would remain on the screen for a randomised duration that lasted between 2 – 10 seconds, which was then followed by a yellow stimulus counter. As soon as the inter-stimulus interval reached the randomised duration, a yellow stimulus counter appeared counting up in milliseconds from 0 – 5 seconds where it would lapse (i.e., error of omission for 0.5 seconds) and begin the next trial, or until the participant tapped on the screen. Once the participant tapped on the screen, their reaction time (i.e., stimulus) would be displayed for 0.5 seconds. At the end of each trial, a black background would appear on-screen for 0.5 seconds. There were 205 trials in total that lasted approximately 25 minutes. Kribbs and Dinges [54] found that after a maximum of three trials, the practice effect for the PVT was removed. This study conservatively implemented five practice trials to ensure participants were fully aware of the task, which were removed from final analyses. If participants responded prematurely during any trial (i.e., before the timer commenced counting up), the trial would reset. To also ensure participants were made aware of their

premature response, the following message in red was displayed on the centre of the screen, ‘You clicked too early! This trial will be reset.’ A visual illustration of the mobile Psychomotor Vigilance Task (m-PVT) is presented in Figure 2.

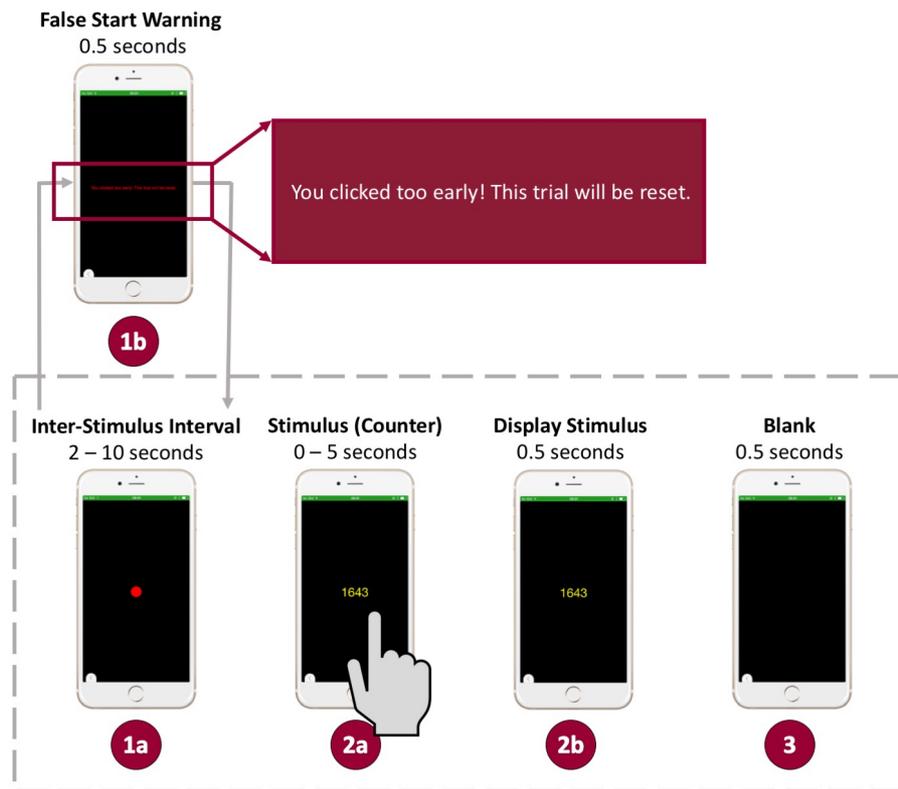


Fig. 1. Mobile Psychomotor Vigilance Task (m-PVT) timeline.

- 1a. Participants were presented with a large red circle (i.e., inter-stimulus interval), which appeared for a randomised duration between 2 – 10 seconds.
- 1b. If participants responded prematurely, a false start warning message appeared informing them that they clicked too early and that the trial would be reset.
- 2a. As soon as the inter-stimulus interval reached the randomised duration, a yellow stimulus counter appeared counting up in milliseconds from 0 – 5 seconds where it would lapse (i.e., error of omission for 0.5 seconds) and begin the next trial, or until the participant had tapped on the screen.
- 2b. Once the participants had tapped on the screen, their reaction time (i.e., stimulus) would be displayed for 0.5 seconds.
3. At the end of each trial, a black background would appear on-screen for 0.5 seconds.



Fig. 2. Visual illustration of the mobile Psychomotor Vigilance Task (m-PVT)

4 Results and Discussion

The aim of the study was to investigate whether an alternative mobile version of the ‘gold standard’ Psychomotor Vigilance Task (PVT) could be used to provide an objective indicator of fatigue in staff working in applied safety critical settings such as train driving, hospital staffs, emergency services, law enforcements, etc. IBM’s Statistical Package for the Social Sciences (SPSS) version 23 for Mac was used to analyse the data. A total of 10,452 test trials were submitted for data analyses, with all 260 practice trials (i.e., 5 practice trials per session) excluded from final analyses. It is important to note that all mobile devices running the online mobile version of the Psychomotor Vigilance Task (m-PVT) were administered through the Dolphin internet browser and were connected using Cardiff University’s Eduroam Wi-Fi roaming service. Therefore, on rare occasions when the Wi-Fi connectivity dropped, the participant’s trial was lost and thus not recorded. As a result, a total of 1.95% ($n = 208$) test trials of all 10,660 trials (i.e., 260 practice and 10,400 test) were lost and not recorded. Based on Basner and Dinges [40] recommendations, all 10,452 test trials with reaction time (RTs) < 100 ms (i.e., false start), which accounted for .05% ($n = 5$) and RTs ≥ 500 ms (i.e., number of lapses), which accounted for 31.84% ($n = 3,328$), were considered for exclusion from the final mean speed response ($1/RT$) and mean reaction time (RT) analyses. All 31.84% ($n = 3,328$) of RTs ≥ 500 ms (i.e., number of lapses) were analysed separately.

4.1 Mean Speed Response (1/RT) and Reaction Time (RT)

Figure 3 presents the illustrated mean speed responses (1/RTs) across the different conditions while Figure 4 presents the illustrated mean reaction times (RTs) across the different conditions. Both the 1/RTs and RTs were submitted to a $2 \times 2 \times 6$ mixed-design analysis of variance (ANOVA) with $2 \times$ mobile devices (iPhone 6s Plus or Samsung Galaxy Tab 4) as the between-subjects factor, and $\times 2$ time of day (Morning, or Afternoon) $\times 6$ time on task (1-minute; 5-minutes, 10-minutes, 15-minutes, 20 minutes, or 25-minutes) as the within-subjects factors. Both the 1/RTs and RTs were significant when comparing the main effect of the two groups using different mobile devices, $F(1, 24)$, 87.21 , $p < .001$, $\eta_p^2 = .78$, indicating a large effect size [55, 56] and $F(1, 24)$, 131.85 , $p < .001$, $\eta_p^2 = .85$, also indicating a large effect size [55, 56], respectively. In addition, there was a significant main effect of time on task for both the 1/RTs and RTs, Wilks' Lambda = .22, $F(5, 20)$, 14.08 , $p < .001$, $\eta_p^2 = .78$, indicating a large effect size [55, 56] and Wilks' Lambda = .24, $F(5, 20)$, 12.66 , $p < .001$, $\eta_p^2 = .76$, indicating a large effect size [55, 56], respectively. Furthermore, there was also a significant interaction between mobile devices \times time on task for both the 1/RTs and RTs, Wilks' Lambda = .34, $F(5, 20)$, 7.95 , $p < .001$, $\eta_p^2 = .67$, indicating a large effect size [55, 56] and Wilks' Lambda = .43, $F(5, 20)$, 5.23 , $p = .003$, $\eta_p^2 = .57$, indicating a moderate effect size [55, 56], respectively. The other main effect (time of day) and interactions (two-way interaction, time of day \times time on task; and three-way interaction, mobile devices \times time of day \times time on task) for both 1/RTs and RTs were not significant.

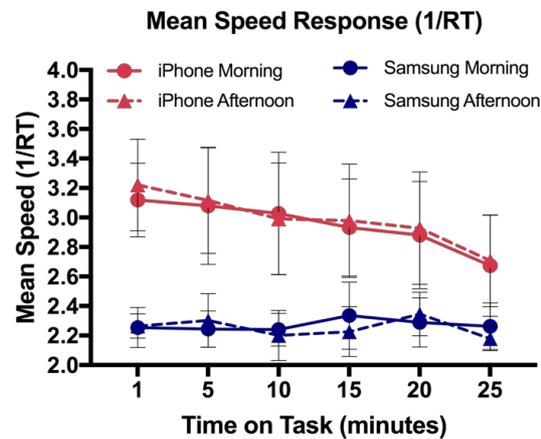


Fig. 3. Mean speed responses (1/RTs) across the different conditions (i.e., morning and afternoon) for both the iPhone 6s Plus and the Samsung Galaxy Tab 4 of the mobile Psychomotor Vigilance Task (m-PVT). *Note:* Mean 1/RTs for both the iPhone 6s Plus and the Samsung Galaxy Tab 4 are presented in bins of 5 minutes as well as the first minute. Error bars represents standard deviation.

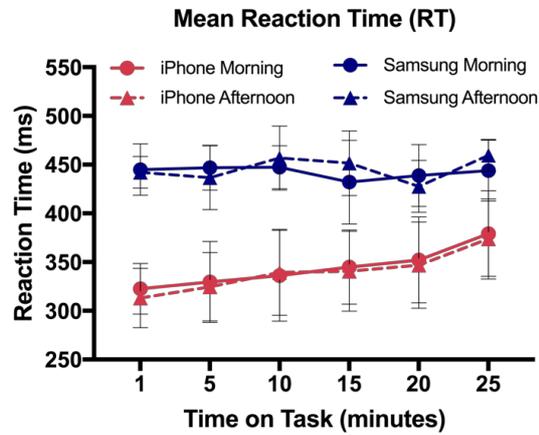


Fig. 4. Mean reaction times (RTs) across the different conditions (i.e., morning and afternoon) for both the iPhone 6s Plus and the Samsung Galaxy Tab 4 of the mobile Psychomotor Vigilance Task (m-PVT). *Note:* Mean RTs for both the iPhone 6s Plus and the Samsung Galaxy Tab 4 are presented in bins of 5 minutes as well as the first minute. Error bars represents standard deviation.

The main effect of the two groups using different mobile devices was followed by post-hoc tests with Bonferroni correction for multiple comparisons. Post-hoc tests showed that participants' mean speed responses ($1/RTs$) were significantly greater with the iPhone 6s Plus mobile device ($M = 2.97$, $SE = .05$) than the Samsung Galaxy Tab 4 mobile device ($M = 2.26$, $SE = .05$, $p < .001$). In addition, post-hoc tests also showed that participants' reaction times (RTs) were significantly faster with the iPhone 6s Plus mobile device ($M = 341.92ms$, $SE = 6.29ms$) than the Samsung Galaxy Tab 4 mobile device ($M = 444.02ms$, $SE = 6.29ms$, $p < .001$). These findings seem to indicate that the iPhone 6s Plus generated significantly greater mean speed responses ($1/RTs$) and significantly faster mean reaction times (RTs) than the Samsung Galaxy Tab 4, with a mean RT difference of 102ms between the iPhone 6s Plus and the Samsung Galaxy Tab 4. Therefore, under these circumstances, the interaction between mobile devices \times time on task was explored separately with a two-way repeated analysis of variance (ANOVA).

iPhone 6s Plus Mean Speed Response ($1/RT$) and Reaction Time (RT)

Figure 5 and Figure 6 present the illustrated mean speed of responses ($1/RTs$) and mean reaction times (RTs) for the iPhone 6s Plus mobile Psychomotor Vigilance Task (m-PVT) across the different conditions. Both the $1/RTs$ and RTs were submitted to a 2×6 two-way repeated analysis of variance (ANOVA) comparing $2 \times$ time of day (Morning, or Afternoon) \times 6 time on task (1-minute; 5-minutes, 10-minutes, 15-minutes, 20 minutes, or 25-minutes). Only the main effect of time on task was significant for both the $1/RTs$ and RTs, Wilks' Lambda = .12, $F(5, 8)$, 12.02, $p = .001$, $\eta_p^2 = .88$, indicating a large effect size [55, 56] and Wilks' Lambda = .12, $F(5, 8)$, 11.93, $p = .002$, $\eta_p^2 = .88$, indicating a large effect size [55, 56], respectively. The other main effect (time of

day) and interactions (two-way interaction, time of day \times time on task) for both 1/RTs and RTs were not significant.

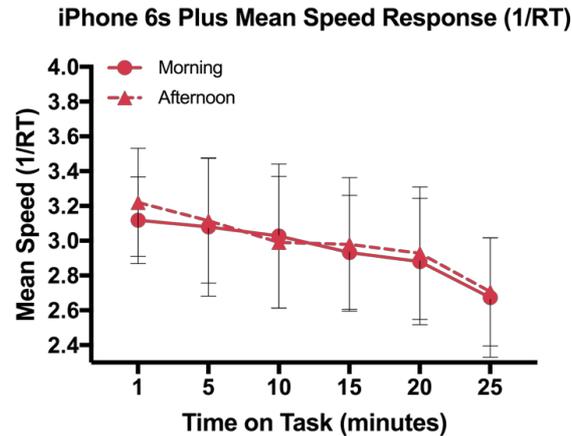


Fig. 5. Mean speed responses (1/RTs) of both the morning session and afternoon session for the iPhone 6s Plus mobile Psychomotor Vigilance Task (m-PVT). *Note:* Mean 1/RTs of both the morning session and afternoon session for the iPhone 6s Plus are presented in bins of 5 minutes as well as the first minute. Error bars represents standard deviation.

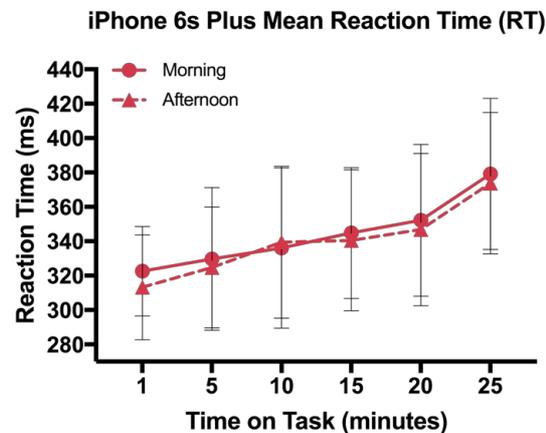


Fig. 6. Mean reaction times (RTs) of both the morning session and afternoon session for the iPhone 6s Plus mobile Psychomotor Vigilance Task (m-PVT). *Note:* Mean RTs of both the morning session and afternoon session are presented in bins of 5 minutes as well as the first minute. Error bars represents standard deviation.

The main effect of time on task was further explored using Fisher's Least Significant Difference (LSD) post-hoc multiple pairwise comparison, which according to Rovai, Baker and Ponton [57] is used when sample sizes are small. As can be seen from Figure 7, participants who were assigned to the iPhone mobile device group had significantly

greater mean speed responses (1/RTs) between the first minute on the m-PVT ($M = 3.17$, $SE = .07$) and 15-minutes on the m-PVT ($M = 2.96$, $SE = .09$, $p = .005$). In addition, participants had significantly greater 1/RTs between the first minute ($M = 3.17$, $SE = .07$) and 20-minutes ($M = 2.90$, $SE = .10$, $p = .005$). Furthermore, participants had significantly greater 1/RTs between the first minute ($M = 3.17$, $SE = .07$) and 25-minutes ($M = 2.69$, $SE = .07$, $p < .001$). Fisher's LSD post-hoc multiple pairwise comparison also showed potential differences between the first minute on the m-PVT ($M = 3.17$, $SE = .07$) and 10-minutes on the m-PVT ($M = 3.01$, $SE = .10$, $p = .051$). However, this was not statistically significant with this study size. As can be seen from Figure 8, participants had significantly faster mean reaction times (RTs) between the first minute on the m-PVT ($M = 317.89$ ms, $SE = 7.09$ ms) and 10-minutes on the m-PVT ($M = 337.75$ ms, $SE = 10.27$ ms, $p = .032$). In addition, participants had significantly faster RTs between the first minute ($M = 317.89$ ms, $SE = 7.09$ ms) and 15-minutes ($M = 342.70$ ms, $SE = 10.22$ ms, $p = .003$). Furthermore, participants had significantly faster RTs between the first minute ($M = 317.89$ ms, $SE = 7.09$ ms) and 20-minutes ($M = 349.52$ ms, $SE = 11.42$ ms, $p = .005$). Moreover, participants had significantly faster RTs between the first minute ($M = 317.89$ ms, $SE = 7.09$ ms) and 25-minutes ($M = 376.47$ ms, $SE = 9.20$ ms, $p < .001$).

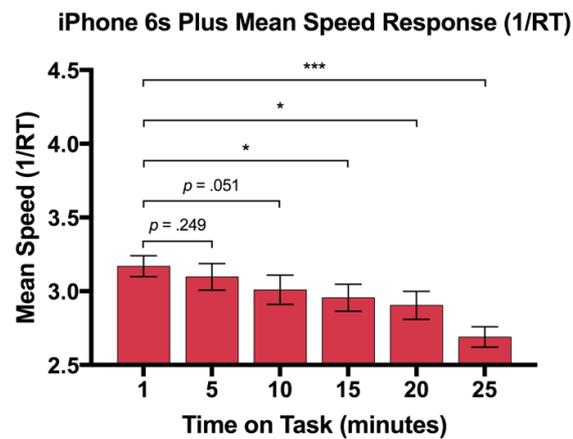


Fig. 7. * $p < .05$; ** $p < .005$; *** $p < .001$. *Note:* Mean speed responses (1/RTs) for the iPhone 6s Plus are presented in bins of 5 minutes as well as the first minute. Error bars represents standard errors.

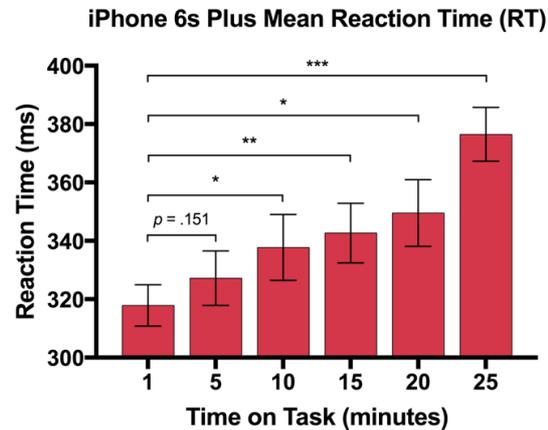


Fig. 8. * $p < .05$; ** $p < .005$; *** $p < .001$. *Note:* Mean reaction times (RTs) for the iPhone 6s Plus are presented in bins of 5 minutes as well as the first minute. Error bars represents standard errors.

Samsung Galaxy Tab 4 Mean Speed Response (1/RT) and Reaction Time (RT)

Figure 9 and Figure 10 present the illustrated mean speed responses (1/RTs) and mean reaction times (RTs) for Samsung Galaxy Tab 4 mobile Psychomotor Vigilance Task (m-PVT) across the different conditions. Both the 1/RTs and RTs were submitted to a 2×6 two-way repeated analysis of variance (ANOVA) comparing $2 \times$ time of day (Morning, or Afternoon) \times 6 time on task (1-minute; 5-minutes, 10-minutes, 15-minutes, 20 minutes, or 25-minutes). For both the 1/RTs and RTs, there was no significant main effect of time of day; Wilks' Lambda = .96, $F(1, 12)$, .530, $p = .481$, $\eta_p^2 = .04$ and Wilks' Lambda = .95, $F(1, 12)$, .579, $p = .461$, $\eta_p^2 = .05$, respectively. In addition, for both the 1/RTs and RTs, there was also no significant main effect of time on task; Wilks' Lambda = .31, $F(5, 8)$, 3.56, $p = .054$, $\eta_p^2 = .69$ and Wilks' Lambda = .31, $F(5, 8)$, 3.53, $p = .056$, $\eta_p^2 = .69$, respectively. Moreover, for both the 1/RTs and RTs, there was also no significant interaction between time of day \times time of task; Wilks' Lambda = .61, $F(5, 8)$, 1.05, $p = .454$, $\eta_p^2 = .40$ and Wilks' Lambda = .63, $F(5, 8)$, .954, $p = .497$, $\eta_p^2 = .37$, respectively.

Samsung Galaxy Tab 4 Mean Speed Response (1/RT)

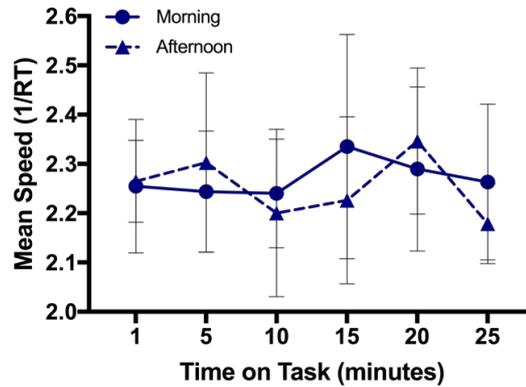


Fig. 9. Mean speed responses (1/RTs) of both the morning session and afternoon session of the Samsung Galaxy Tab 4 mobile Psychomotor Vigilance Task (m-PVT). *Note:* Mean 1/RTs of both the morning session and afternoon session for the Samsung Galaxy Tab 4 are presented in bins of 5 minutes as well as the first minute. Error bars represents standard deviation.

Samsung Galaxy Tab 4 Mean Reaction Time (RT)

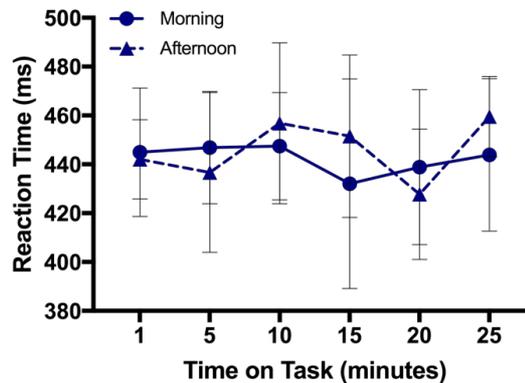


Fig. 10. Mean reaction times (RTs) of both the morning session and afternoon session of the Samsung Galaxy Tab 4 mobile Psychomotor Vigilance Task (m-PVT). *Note:* Mean RTs of both the morning session and afternoon session for the Samsung Galaxy Tab 4 are presented in bins of 5 minutes as well as the first minute. Error bars represents standard deviation.

4.2 Mean Number of Lapses

From all test trials, a total of 31.84% ($n = 3,328$) RTs ≥ 500 ms were submitted for data analyses. Figure 11 presents the illustrated mean number of lapses across the different conditions. The mean number of lapses were submitted to a $2 \times 2 \times 6$ mixed-design analysis of variance (ANOVA) with $2 \times$ mobile devices (iPhone 6s Plus or Samsung

Galaxy Tab 4) as the between-subjects factor, and $\times 2$ time of day (Morning, or Afternoon) $\times 6$ time on task (1-minute; 5-minutes, 10-minutes, 15-minutes, 20 minutes, or 25-minutes) as the within-subjects factors. There was a significant main effect of the two groups using different mobile devices, $F(1, 24)$, 131.81, $p < .001$, $\eta_p^2 = .85$, indicating a large effect size [55, 56]. In addition, there was a significant main effect of time on task, Wilks' Lambda = .28, $F(5, 20)$, 10.27, $p < .001$, $\eta_p^2 = .72$, indicating a large effect size [55, 56]. Furthermore, there was also a significant interaction between mobile devices \times time on task, Wilks' Lambda = .31, $F(5, 20)$, 9.10, $p < .001$, $\eta_p^2 = .70$, indicating a large effect size [55, 56]. The other main effect (time of day, $p = .620$) and interactions (two-way interaction, time of day \times time on task, $p = .395$; and three-way interaction, mobile devices \times time of day \times time on task, $p = .151$) for the mean number of lapses (i.e., RTs ≥ 500 ms) were not significant.

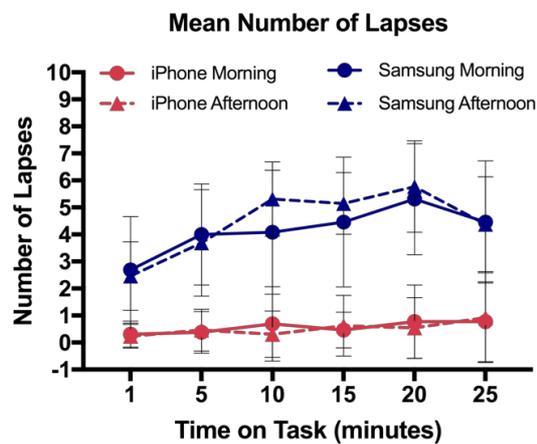


Fig. 11. Mean number of lapses across the different conditions (i.e., morning and afternoon) for both the iPhone 6s Plus and the Samsung Galaxy Tab 4 of the mobile Psychomotor Vigilance Task (m-PVT). *Note:* Mean number of lapses for both the iPhone 6s Plus and the Samsung Galaxy Tab 4 are presented in bins of 5 minutes as well as the first minute. Error bars represent standard deviation.

The main effect of the two groups using different mobile devices was followed by post-hoc tests with Bonferroni correction for multiple comparisons. Post-hoc tests showed that participants' mean number of lapses were significantly lower for the iPhone 6s Plus mobile device ($M = .54$, $SE = .23$) than the Samsung Galaxy Tab 4 mobile device ($M = 4.31$, $SE = .23$, $p < .001$). These findings seem to indicate that participants assigned to the iPhone 6s Plus recorded significantly less mean number of lapses than the Samsung Galaxy Tab 4. These findings are not too surprising as it was previously found that both the mean speed responses ($1/RTs$) and mean reaction times (RTs) for the iPhone 6s Plus generated significantly greater $1/RTs$ and faster RTs than the Samsung Galaxy Tab 4. There was a statistically difference of 102ms, which would indicate at least for the Samsung Galaxy Tab 4 that there would be significantly more test trials with RTs ≥ 500 ms (i.e., number of lapses). As a result, from all 31.84% (n

= 3,328) of test trials with RTs \geq 500 ms, the Samsung Galaxy Tab 4 group represented 90.32% ($n = 3,006$) and the iPhone 6s Plus group represented 9.68% ($n = 322$). Therefore, also under these circumstances, the interaction between mobile devices \times time on task was explored separately with a two-way repeated analysis of variance (ANOVA).

iPhone 6s Plus Mean Number of Lapses

Figure 12 presents the illustrated mean number of lapses for the iPhone 6s Plus mobile Psychomotor Vigilance Task (m-PVT) across the different conditions. The mean number of lapses were submitted to a 2×6 two-way repeated analysis of variance (ANOVA) comparing $2 \times$ time of day (Morning, or Afternoon) \times 6 time on task (1-minute; 5-minutes, 10-minutes, 15-minutes, 20 minutes, or 25-minutes). There was no significant main effect of time of day; Wilks' Lambda = .997, $F(1, 12)$, .04, $p = .846$, $\eta_p^2 = .00$. In addition, there was also no significant main effect of time on task; Wilks' Lambda = .75, $F(5, 8)$, .54, $p = .744$, $\eta_p^2 = .25$. Moreover, there was also no significant interaction between time of task \times time of day; Wilks' Lambda = .36, $F(5, 8)$, 2.84, $p = .092$, $\eta_p^2 = .64$.

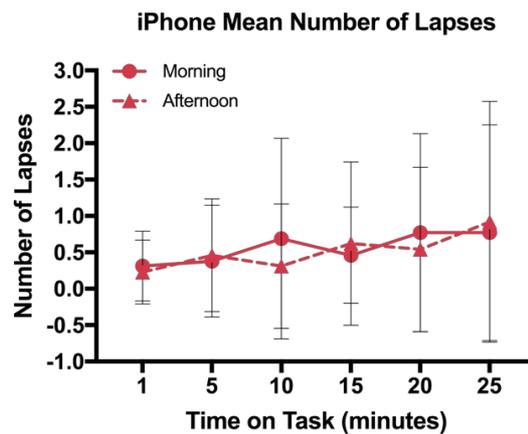


Fig. 12. Mean number of lapses for both the morning session and afternoon session for the iPhone 6s Plus of the mobile Psychomotor Vigilance Task (m-PVT). *Note:* Mean number of lapses for the iPhone 6s Plus are presented in bins of 5 minutes as well as the first minute. Error bars represents standard deviation.

Samsung Galaxy Tab 4 Mean Number of Lapses

Figure 13 presents the illustrated mean number of lapses for the Samsung Galaxy Tab 4 mobile Psychomotor Vigilance Task (m-PVT) across the different conditions. The mean number of lapses were submitted to a 2×6 two-way repeated analysis of variance (ANOVA) comparing $2 \times$ time of day (Morning, or Afternoon) \times 6 time on task (1-minute; 5-minutes, 10-minutes, 15-minutes, 20 minutes, or 25-minutes). Only the main effect of time on task was significant for the mean number of lapses, Wilks' Lambda = .14, $F(5, 8)$, 9.80, $p = .003$, $\eta_p^2 = .86$, indicating a large effect size [55, 56]. The other

main effect (time of day, $p = .486$) and two-way interaction (time of day \times time on task, $p = .227$) for the mean number of lapses (i.e., RTs ≥ 500 ms) were not significant.

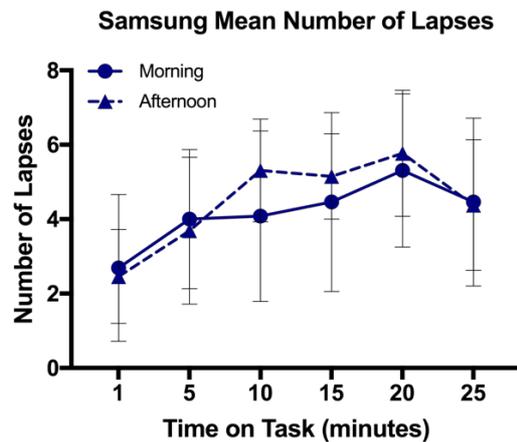


Fig. 13. Mean number of lapses for both the morning session and afternoon session for the Samsung Galaxy Tab 4 of the mobile Psychomotor Vigilance Task (m-PVT). *Note:* Mean number of lapses for the Samsung Galaxy Tab 4 are presented in bins of 5 minutes as well as the first minute. Error bars represents standard deviation.

The main effect of time on task was further explored using Fisher's Least Significant Difference (LSD) post-hoc multiple pairwise comparison, which according to Rovai, Baker and Ponton [57] is used when sample sizes are small. As can be seen from Figure 14, participants who were assigned to the Samsung Galaxy Tab 4 mobile device group had significantly less mean number of lapses between the first minute on the m-PVT ($M = 2.58$, $SE = .35$) and 5-minutes on the m-PVT ($M = 3.85$, $SE = .37$, $p = .001$). In addition, participants also had significantly less mean number of lapses between the first minute on the m-PVT ($M = 2.58$, $SE = .35$) and 10-minutes on the m-PVT ($M = 4.69$, $SE = .40$, $p < .001$). Furthermore, participants also had significantly less mean number of lapses between the first minute on the m-PVT ($M = 2.58$, $SE = .35$) and 15-minutes on the m-PVT ($M = 4.81$, $SE = .40$, $p = .001$). Moreover, participants also had a significantly lower mean number of lapses between the first minute on the m-PVT ($M = 2.58$, $SE = .35$) and 20-minutes on the m-PVT ($M = 5.54$, $SE = .38$, $p < .001$). Finally, participants also had a significantly lower mean number of lapses between the first minute on the m-PVT ($M = 2.58$, $SE = .35$) and 25-minutes on the m-PVT ($M = 4.42$, $SE = .46$, $p = .008$). These findings seem to indicate that mean number of lapses for mobile devices, that generate on average significantly slower thresholds, due to perhaps hardware configurations than what is typically found in the Psychomotor Vigilance Task (PVT) literature, may not be an accurate representation and comparison from analyses of both the mean speed responses ($1/RT$) and mean reaction times (RTs). Instead, the analyses of the mean number of lapses may yield far better research insights.

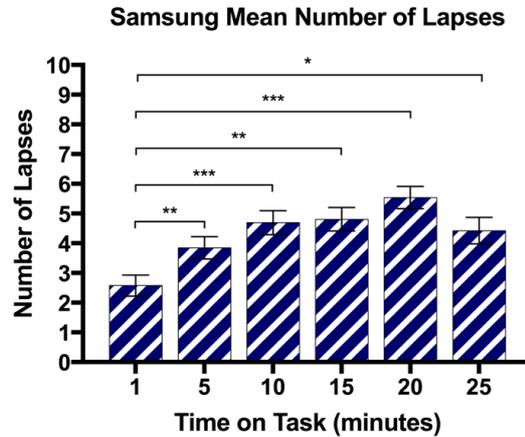


Fig. 14. * $p < .05$; ** $p < .005$; *** $p < .001$. *Note:* Mean number of lapses for the Samsung Galaxy Tab 4 are presented in bins of 5 minutes as well as the first minute. Error bars represents standard errors.

5 Conclusion

The study aimed to investigate whether an alternative online mobile version of the ‘gold standard’ Psychomotor Vigilance Task (PVT) could be used to provide an objective indicator of fatigue in staff in applied safety critical settings such as train driving, hospital staffs, emergency services, law enforcements, etc. It was found that there was a large significant difference in reaction times (RTs) between the two mobile devices (i.e., Samsung vs. Apple’s iPhone). Apple’s iPhone 6s Plus generated RTs that were comparable to those found in the literature [33, 34, 40, 42–46, 52]. However, the RTs of the Samsung mobile device were significantly slower than those found in the literature. Findings from this study also support previous research that have identified that an increase in fatigue results in impaired alertness [58, 59], whereby sustained attention, as measured by reaction time, significantly reduces after 10-minutes of continuous performance using the Psychomotor Vigilance Task (PVT). These findings from this alternative online mobile version of the Psychomotor Vigilance Task (m-PVT) are consistent with previous work, which suggested that sustained attention drops with prolonged duration of the task [60, 61].

This study seems to suggest that an alternative online mobile version of the ‘gold standard’ 10-min PVT (i.e., m-PVT) could be used to provide an objective indicator of fatigue after 10 minutes on the m-PVT in staff working in applied safety critical settings such as train driving, hospital staffs, emergency services, law enforcements, etc. However, caution is required when considering implementing an alternative online mobile version (m-PVT) that is running on an internet browser, as only the iPhone 6s Plus was able to generate reaction times that were comparable with the literature. In contrast, there were significantly fewer lapses for the iPhone 6s Plus ($n = 322$) than the Samsung Galaxy Tab 4 ($n = 3,006$), which was not surprising when considering that both mean

speed responses (1/RTs) and reaction times (RTs) were significantly higher and faster respectively, for the iPhone 6s Plus than for the Samsung Galaxy Tab 4. As a result, perhaps analyses of both the mean speed responses (1/RTs) and mean reaction times (RTs) may not always generate an accurate data representation for analyses based on the hardware differences in mobile manufactures as well as configurations and specifications. Therefore, perhaps using the number of lapses (i.e., RTs \geq 500 ms) may yield richer data for analyses on these circumstances. As a result, this study recommends that pilot studies should be carried out to firstly explore and determine whether the selected mobile device generates RTs that are better suited for either mean RTs and mean 1/RTs, or mean number of lapses analyses. However, there are several factors that could also account for the difference in the mean 1/RTs and mean RTs between the two mobile devices.

Firstly, regarding software, both the Apple's iPhone 6s Plus and Samsung Galaxy Tab 4 run on different operating systems (OS). Apple's iPhone 6s Plus run their own native iOS version 9.3.1, while the Samsung Galaxy Tab 4 run on Alphabet's Android KitKat version 4.4.2. Furthermore, even though the same internet browser (Dolphin Web Browser) was used across both mobile devices, the version numbers were different. This may indicate that one may have had more improvement and stability updates than the other (Dolphin Web Browser; Apple's native iOS app version 9.9.0 vs. Android app OS version 11.5.6). Alternatively, the browser may have been developed for one platform and then expanded to also run on the other platform.

5.1 Future Work

Further research is now needed to determine whether the m-PVT can be used to provide an objective indicator of fatigue in staff working in applied safety critical settings such as train driving, hospital staffs, emergency services, law enforcement, etc. Use of an iPhone 6s Plus is recommended, and further studies with larger samples are required to confirm the length of the task.

6 Acknowledgment

This work was supported by the Economic and Social Research Council [ES/J500197/1] and Arriva Trains Wales (ATW).

7 Contributors

MSE designed the experiment, carried out data collection, statistical analysis and writing of the paper. DH programmed the Psychomotor Vigilance Task (m-PVT) under the directions of MSE using the client code HTML, and CSS for the page visualisation and layout. DH also used the JavaScript to initiate the m-PVT. MSE wrote the paper with input from APS. All authors have seen and approved the final version of the paper for publication.

8 References

1. Eggemeier, F.T., Wilson, G.F., Kramer, A.F., Damos, D.L.: General considerations concerning workload assessment in multi-task environments. In: Damos, D. L. (eds.) *Multiple Task Performance*. pp. 207–216. Taylor & Francis, London 1991.
2. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (task load index): Results of empirical and theoretical research. *Advances in Psychology*. **52**, 139–183 1988. doi: 10.1016/S0166-4115(08)62386-9
3. Longo, L.: Designing medical interactive systems via assessment of human mental workload. In: *International Symposium on Computer-Based Medical Systems*. pp. 364–365 2015. doi: 10.1109/CBMS.2015.67
4. Longo, L.: Subjective Usability, Mental Workload Assessments and Their Impact on Objective Human Performance. In: *IFIP Conference on Human-Computer Interaction*. pp. 202–223. 2017. doi: 10.1007/978-3-319-67684-5_13
5. Zammouri, A., Moussa, A.A., Mebrouk, Y.: Brain-computer interface for workload estimation: Assessment of mental efforts in learning processes. *Expert Systems with Applications*. **112**, 138–147 2018. doi: 10.1016/j.eswa.2018.06.027
6. Byrne, A.J., Oliver, M., Bodger, O., Barnett, W.A., Williams, D., Jones, H., Murphy, A.: Novel method of measuring the mental workload of anaesthetists during clinical practice. *British Journal of Anaesthesia*. **105**, 767–771 2010. doi: 10.1093/bja/aeq240
7. Orlandi, L., Brooks, B.: Measuring mental workload and physiological reactions in marine pilots: Building bridges towards redlines of performance. *Applied Ergonomics*. **69**, 74–92 2018. doi: 10.1016/j.apergo.2018.01.005
8. Shakouri, M., Ikuma, L.H., Aghazadeh, F., Nahmens, I.: Analysis of the sensitivity of heart rate variability and subjective workload measures in a driving simulator: The case of highway work zones. *International Journal of Industrial Ergonomics*. **66**, 136–145 2018. doi: 10.1016/j.ergon.2018.02.015
9. Comstock, J.R.J., Arnegard, R.J.: *The Multi-Attribute Test Battery for Human Operator Workload and Strategic Behaviour*. National Aeronautics and Space Administration, Hampton 1992.
10. Berka, C., Levendowski, D.J., Lumicao, M.N., Yau, A., Davis, G., Zivkovic, V.T., Olmstead, R.E., Tremoulet, P.D., Craven, P.L.: EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation Space and Environmental Medicine*. **78**, B231–B244 2007.
11. Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., Babiloni, F.: Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience and Biobehavioral Reviews*. **44**, 58–75 2014. doi: 10.1016/j.neubiorev.2012.10.003
12. Heine, T., Lenis, G., Reichensperger, P., Beran, T., Doessel, O., Deml, B.: Electrocardiographic features for the measurement of drivers' mental workload. *Applied Ergonomics*. **61**, 31–43. 2017. doi: 10.1016/j.apergo.2016.12.015
13. Hogervorst, M.A., Brouwer, A.M., Van Erp, J.B.E.: Combining and comparing EEG, peripheral physiology and eye-related measures for the assessment of mental workload. *Frontiers in Neuroscience*. **8**, 322 2014. doi: 10.3389/fnins.2014.00322
14. Hsu, B.W., Wang, M.J.J., Chen, C.Y., Chen, F.: Effective indices for monitoring mental workload while performing multiple tasks. *Perceptual and Motor Skills*. **121**, 94–117 2015. doi: 10.2466/22.PMS.121c12x5
15. Jimenez-Molina, A., Retamal, C., Lira, H.: Using psychophysiological sensors to assess mental workload during web browsing. *Sensors*. **18**, 458 2018. doi: 10.3390/s18020458

16. Shaw, E.P., Rietschel, J.C., Hendershot, B.D., Pruziner, A.L., Miller, M.W., Hatfield, B.D., Gentili, R.J.: Measurement of attentional reserve and mental effort for cognitive workload assessment under various task demands during dual-task walking. *Biological Psychology*. **134**, 39–51 2018. doi: 10.1016/j.biopsycho.2018.01.009
17. So, W.K.Y., Wong, S.W.H., Mak, J.N., Chan, R.H.M.: An evaluation of mental workload with frontal EEG. *PLOS ONE*. **12**, e0174949 2017. doi: 10.1371/journal.pone.0174949
18. Vergara, R.C., Moenne-Loccoz, C., Maldonado, P.E.: Cold-Blooded attention: Finger temperature predicts attentional performance. *Frontiers in Human Neuroscience*. **11**, 454 2017. doi: 10.3389/fnhum.2017.00454
19. Widyanti, A., Hanna, Muslim, K., Sitalaksana, I.Z.: The sensitivity of Galvanic Skin Response for assessing mental workload in Indonesia. *Work: A Journal of Prevention Assessment & Rehabilitation*. **56**, 111–117 2017. doi: 10.3233/WOR-162479
20. Blanco, J.A., Johnson, M.K., Jaquess, K.J., Oh, H., Lo, L.C., Gentili, R.J., Hatfield, B.D.: Quantifying Cognitive Workload in Simulated Flight Using Passive, Dry EEG Measurements. *IEEE Transactions on Cognitive and Developmental Systems*. **10**, 373–383 2018. doi: 10.1109/TCDS.2016.2628702
21. Myrtek, M., Deutschmannjanicke, E., Strohmaier, H., Zimmermann, W., Lawerenz, S., Brugner, G., Muller, W.: Physical, mental, emotional, and subjective workload components in train drivers. *Ergonomics*. **37**, 1195–1203 1994. doi: 10.1080/00140139408964897
22. Brookhuis, K.A., De Waard, D.: Assessment of drivers' workload: Performance and subjective and physiological indexes. In: Hancock, P.A., Desmond, P.A. (eds.) *Stress, Workload and Fatigue*. pp. 321–333. Lawrence Erlbaum Associates, New Jersey 2001.
23. Foy, H.J., Chapman, P.: Mental workload is reflected in driver behaviour, physiology, eye movements and prefrontal cortex activation. *Applied Ergonomic*. **73**, 90–99 2018. doi: 10.1016/j.apergo.2018.06.006
24. Paxion, J., Galy, E., Berthelon, C.: (2014). Mental workload and driving. *Frontiers in Psychology*. **5**, 1344 2014. doi: 10.3389/fpsyg.2014.01344
25. Wickens, C.D.: Mental workload: assessment, prediction and consequences. In *International Symposium on Human Mental Workload: Models and Applications*. pp. 18–29. Springer, Cham 2017. doi: 10.1007/978-3-319-61061-0_2
26. Gopher, D., Donchin, E.: Workload: An examination of the concept. In: Boff, K.R., Kaufman, L., Thomas, J.P. (eds.) *Handbook of perception and human performance*, Vol. 2. Cognitive processes and performance. pp. 1–49. John Wiley & Sons, Oxford, England 1986.
27. Hancock, P.A., Meshkati, N.: *Human Mental Workload*. Elsevier Science, Amsterdam 1988.
28. Moray, N.: *Mental Workload: Its Theory and Measurement*. Plenum, New York 1979.
29. Xie, B., Salvendy, G.: Review and reappraisal of modelling and predicting mental workload in single- and multi-task environments. *Work and Stress*. **14**, 74–99 2000. doi: 10.1080/026783700417249
30. Byrne, A.: Mental Workload as an Outcome in Medical Education. In *International Symposium on Human Mental Workload: Models and Applications*. pp. 187–197. Springer, Cham. 2017. doi: doi.org/10.1007/978-3-319-61061-0_12
31. Gaba, D.M., & Lee, T.: (1990). Measuring the workload of the anesthesiologist. *Anesthesia and Analgesia*. **71**, 354–361 1990.
32. Smith, A. P., Smith, H.: Workload, fatigue and performance in the rail industry. In *International Symposium on Human Mental Workload: Models and Applications*. pp. 251–263. 2017. Springer, Cham. doi: 10.1007/978-3-319-61061-0_17
33. Dinges, D.F., Powell, J. W.: Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers*. **17**, 652–655 (1985). doi: 10.3758/BF03200977

34. Dinges, D.F., Orne, M.T., Whitehouse, W.G., Orne, E.C.: Temporal placement of a nap for alertness: contributions of circadian phase and prior wakefulness. *Sleep*. **10**, 313–329 (1987).
35. Belenky, G., Wesensten, N.J., Thorne, D.R., Thomas, M.L., Sing, H.C., Redmond, D.P., Russo, M.B., Balkin, T.J.: Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. *Journal of Sleep Research*. **12**, 1–12 (2003). doi: 10.1046/j.1365-2869.2003.00337.x
36. Dinges, D.F., Pack, F., Williams, K., Gillen, K.A., Powell, J.W., Ott, G.E., Aptowicz, C., Pack, A.I.: Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. *Sleep*. **20**, 267–777 (1997).
37. Jewett, M.E., Dijk, D.J., Kronauer, R.E., Dinges, D.F.: Dose-response relationship between sleep duration and human psychomotor vigilance and subjective alertness. *Sleep*. **22**, 171–179 (1999). doi: 10.1093/sleep/22.2.171
38. Lamond, N., Dorrian, J., Roach, G.D., McCulloch, K., Holmes, A.L., Burgess, H.J., Fletcher, A., Dawson, D.: The impact of a week of simulated night work on sleep, circadian phase, and performance. *Occupational and Environmental Medicine*. **60**, 1–9 (2003). doi: 10.1136/oem.60.11.e13
39. Davis, C.M., Roma, P.G., Hienz, R.D.: A rodent model of the human psychomotor vigilance test: Performance comparisons. *Journal of Neuroscience Methods*. **259**, 57–71 (2016). doi: 10.1016/j.jneumeth.2015.11.014
40. Basner, M., Dinges, D.F.: Maximizing sensitivity of the Psychomotor Vigilance Test (PVT) to sleep loss. *Sleep*. **34**, 581–591 (2011). doi: 10.1093/sleep/34.5.581
41. Van Dongen, H.P., Dinges, D.F.: Sleep, circadian rhythms, and psychomotor vigilance. *Clinics in Sports Medicine*. **24**, 237–249 (2005). doi: 10.1016/j.csm.2004.12.007
42. Basner, M., Mollicone, D.J., Dinges, D.F.: Validity and sensitivity of a brief Psychomotor Vigilance Test (PVT-B) to total and partial sleep deprivation. *Acta Astronautica*. **69**, 949–959 (2011). doi: 10.1016/j.actaastro.2011.07.015
43. Loh, S., Lamond, N., Dorrian, J., Roach, G., Dawson, D.: The validity of psychomotor vigilance tasks of less than 10-minute duration. *Behavior Research Methods, Instruments, & Computers*. **36**, 339–346 (2004). doi: 10.3758/BF03195580
44. Lamond, N., Dawson, D., Roach, G.D.: Fatigue assessment in the field: Validation of a hand-held electronic psychomotor vigilance task. *Aviation, Space, and Environmental Medicine*. **76**, 486–489 (2005).
45. Roach, G.D., Dawson, D., Lamond, N.: Can a shorter psychomotor vigilance task be used as a reasonable substitute for the ten-minute psychomotor vigilance task? *Chronobiology International*. **23**, 379–387 (2006). doi: 10.1080/07420520601067931
46. Dorrian, J., Roach, G.D., Fletcher, A., Dawson, D.: Simulated train driving: Fatigue, self-awareness and cognitive disengagement. *Applied Ergonomics*. **38**, 155–166 (2007). doi: 10.1016/j.apergo.2006.03.006
47. Lamond, N., Jay, S.M., Dorrian, J., Ferguson, S.A., Roach, G.D., Dawson, D.: The sensitivity of a palm-based psychomotor vigilance task to severe sleep loss. *Behavior Research Methods*. **40**, 347–352 (2008). doi: 10.3758/BRM.40.1.347
48. Dorrian, J., Rogers, N.L., Dinges, D.F.: Psychomotor vigilance performance: Neurocognitive assay sensitive to sleep loss. In: Kushida, C.A. (eds.) *Sleep deprivation: clinical Issues. pharmacology and sleep loss effects*. pp. 39–70. Marcel Dekker Inc., New York (2005).
49. Basner, M., Mcguire, S., Goel, N., Rao, H., Dinges, D.F.: A new likelihood ratio metric for the psychomotor vigilance test and its sensitivity to sleep loss. *Journal of Sleep Research*. **24**, 702–713 (2015). doi: 10.1111/jsr.12322

50. Dinges, D.F., Kribbs, N.B.: Performing while sleepy: Effects of experimentally-induced sleepiness. In: Monk, T.H. (eds.) *Sleep, Sleepiness and Performance*. pp. 97–128. John Wiley and Sons, Ltd., Chichester (1991).
51. Warm, J.S., Parasuraman, R., Matthews, G.: Vigilance requires hard mental work and is stressful. *Human Factors*. **50**, 43–441 (2008). doi: 10.1518/001872008X312152
52. Khitrov, M.Y., Laxminarayan, S., Thorsley, D., Ramakrishnan, S., Rajaraman, S., Wesensten, N.J., Reifman, J.: PC-PVT: A platform for psychomotor vigilance task testing, analysis, and prediction. *Behavior Research Methods*. **46**, 140–147 (2014). doi: 10.3758/s13428-013-0339-9
53. World Medical Association: World Medical Association Declaration of Helsinki Ethical Principles for Medical Research Involving Human Subjects. *Jama-Journal of the American Medical Association*. **310**, 2191–2194 (2013). doi: 10.1001/jama.2013.281053
54. Kribbs, N.B., Dinges, D.F.: Vigilance decrement and sleepiness. In: Ogilvie, R.D., Harsh, J. (eds.) *Sleep Onset: Normal and abnormal processes*. pp. 113–125. American Psychological Association, Washington, D.C. (1994).
55. Cohen, J.: Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement*. **33**, 107–112 (1973). doi: 10.1177/001316447303300111
56. Cohen, J.: *Statistical power analysis for the behavioral sciences*. Erlbaum, Hillsdale, NJ (1988).
57. Rovai, A.P., Baker, J.D., Ponton, M.K.: *Social science research design and statistics: A practitioner’s guide to research methods and IMB SPS analysis*. Watertree Press LLC, Chesapeake, VA (2014).
58. Dorrian, J., Hussey, F., Dawson, D.: Train driving efficiency and safety: Examining the cost of fatigue. *Journal of Sleep Research*. **16**, 1–11 (2007). doi: 10.1111/j.1365-2869.2007.00563.x
59. Dorrian, J., Baulk, S.D., Dawson, D.: Work hours, workload, sleep and fatigue in Australian Rail Industry employees. *Applied Ergonomics*. **42**, 202–209 (2011). doi: 10.1016/j.apergo.2010.06.009
60. Dinges, D.F., Powell, J.W.: Sleepiness is more than lapsing. *Journal of Sleep Research*. **17**, 84 (1988).
61. Dinges, D.F., Powell, J.W.: Sleepiness impairs optimum response capability: It’s time to move beyond the lapse hypothesis. *Journal of Sleep Research*. **18**, 366 (1989).