

Preliminary validation study of the 3-min wrist-worn psychomotor vigilance test

Panagiotis Matsangas¹ · Nita Lewis Shattuck¹ · Stephanie Brown¹

Published online: 28 October 2016
© Psychonomic Society, Inc. (outside the USA) 2016

Abstract The 10-min psychomotor vigilance test (PVT) is a widely used tool to assess behavioral alertness. In operational environments, however, the 3-min version of the PVT is more practicable. Under controlled laboratory conditions, we compared performance on the 3-min PVT on a wrist-worn device with the 3-min PVT on a laptop computer. We conducted two repeated measures experiments (72 participants) to assess the effects of PVT device type (laptop or wrist-worn), backlight on the wrist-worn device (on or off), ambient light (low or normal levels), and interstimulus interval (ISI). First, we compared the wrist-worn PVT (ISI of 2–10 s) with the laptop-based PVT (ISI of 1–4 s). Participants performed six PVT trials counterbalanced for order of ambient light and device type. In both ambient-lighting conditions, the median differences in PVT metrics (reaction time, response speed, and percentage of 355-ms or 500-ms lapses combined with false starts) between the laptop and the wrist-worn PVT with the backlight on were less than 4.5 %. Specifically, the median difference in reaction times was less than 10 ms. However, the standard deviation in most metrics using the wrist-worn PVT was twice that of the laptop. In a second experiment, we compared the wrist-worn and laptop PVTs with the same ISIs (1–4 s) in normal light conditions. The results suggest that PVT data can be collected reliably in the field using the wrist-worn device with the backlight on and an ISI of 2–10 s.

Keywords Psychomotor vigilance task · Reaction time · Field data collection

Many tasks in military operations include a vigilance component requiring the operator to pay attention for prolonged periods of time to detect the infrequent occurrence of critical events. It is well-documented that the human ability to sustain focused attention deteriorates over time. This phenomenon is further exacerbated by sleepiness and fatigue, both of which are common problems in military and nonmilitary operational environments. One task used widely in laboratory and field studies to assess sustained attention is the visual psychomotor vigilance task (PVT; Dinges & Powell, 1985). The PVT is a simple reaction time task in which participants are required to press a response button as soon as a stimulus appears on a display screen. PVT performance not only is affected by sleep loss but is also sensitive to circadian rhythmicity (Dinges et al., 1997; Doran, Van Dongen, & Dinges, 2001; Durmer & Dinges, 2005; Jewett, Dijk, Kronauer, & Dinges, 1999; Wyatt et al., 1997). Because of its simplicity, the PVT has only minor learning effects. Asymptotic performance can be reached in one to three trials (Balkin et al., 2000; Dinges et al., 1997; Jewett et al., 1999; Kribbs & Dinges, 1994; Rosekind et al., 1994).

The typical PVT duration is 10 min, with an interstimulus interval (ISI) of 2–10 s. Even though it is used extensively in laboratory settings, the 10-min version of the traditional PVT is problematic in operational field studies for a variety of reasons. For one thing, study participants often refuse to be pulled away from their actual work to engage in an artificial 10-min task. In addition, providing a laptop to each individual participant in a study is not cost-effective, and, due to security concerns, some operational environments will not allow for

✉ Panagiotis Matsangas
pmatsang@nps.edu

¹ Operations Research Department, Naval Postgraduate School, 1411 Cunningham Drive, Monterey, CA 93943, USA

the introduction of laptops or other devices that could potentially transmit information. Centrally locating the testing device in a library or a common access area requires participants to go out of their way to participate—again adding to the requirements of an already overbooked schedule. In light of these problems, it is no surprise that both researchers (Lamond et al., 2008) and crewmembers on U.S. Navy ships consider the original 10-min PVT excessively long for field studies.

An alternative to the original PC-based PVT is a version that is embedded in hand-held devices like a smart phone or personal digital assistant (Rosekind, Gregory, & Mallis, 2006). Research has shown that PVT versions of 3–5 min duration provide results comparable to those from the original 10-minute PVT, and can be used to reliably detect effects of sleep loss (Basner, Mollicone, & Dinges, 2011; Lamond, Dawson, & Roach, 2005; Thorne et al., 2005) and fatigue-related performance decrements (Basner & Rubinstein, 2011). However, one study has reported that the shorter 5-min PVT may be less sensitive to sleep loss effects than the original 10-min version (Lamond et al., 2008).

For more than 10 years, researchers at the Naval Postgraduate School have conducted numerous studies using actigraphy to assess sleep patterns in the military operational domain (Miller, Matsangas, & Kenney, 2012). In more recent efforts, we have used a type of actiwatch that includes an embedded version of the PVT to assess the effect of shiftwork on psychomotor vigilance performance (see, e.g., Shattuck, Matsangas, & Brown, 2015; Shattuck, Matsangas, & Powley, 2015; Shattuck, Matsangas, & Waggoner, 2014; Shattuck, Waggoner, Young, Smith, & Matsangas, 2014). While wearing the wrist-worn device (actiwatch) to assess activity and sleep patterns, participants can also use the same wrist-worn device to take the PVT without having to leave their workplace or carry additional equipment. The use of the PVT integrated with the wrist-worn device yielded good results. Participants' reported that the wrist-worn PVT was easy to use. Additionally, participant compliance when using the wrist-worn version was much higher, and the attrition rate was reduced as compared to the PC-based PVT that had been used in our earlier studies.

Although it has proven useful, the wrist-worn PVT device is fairly novel; its use could potentially introduce additional confounding factors that could affect PVT performance metrics. The different interface design, the screen characteristics, and the ambient-lighting conditions are some of the factors that may have differential effects on PVT performance in the two devices. Our review of the literature failed to identify any studies comparing the 3-min PVT embedded in the wrist-worn device with the computer-based PVT. Given these concerns, this study had two objectives. The main objective was to compare the results from the wrist-worn 3-min PVT with an ISI of 2–10 s with those from the validated computer-based 3-min

PVT with an ISI of 1–4 s. The second objective was to assess two potentially confounding issues with the wrist-worn device—specifically, the effect of the backlight feature provided in the wrist-worn device, and the effect of ambient lighting.

Method

Participants

Seventy-two individuals, on average 34.6 ± 7.70 years of age, from the Naval Postgraduate School (NPS) volunteered to participate in the study. The participants were screened for corrected vision, recent injuries or pain in the arms, wrists, or fingers, or a diagnosis of color vision deficiency or carpal tunnel syndrome. The study protocol was approved by the NPS Institutional Review Board. Informed consent was obtained after the experimental procedures had been fully explained.

Equipment

A study questionnaire was developed that included demographic questions, sleep history for the 48 h prior to the data collection, the current day's caffeine intake, and any issues that might affect vision. Psychomotor vigilance performance data were collected with two devices: a laptop with a validated version of the PVT (PULSAR Informatics, Philadelphia, PA), and a wrist-worn device (Motionlogger Watch) with an embedded version of the PVT to be validated (Ambulatory Monitoring, Inc., Ardsley, NY). The wrist-worn device uses the Infrared Data Association (IrDA) wireless optical communication protocol to transmit data up to 1 m in range.

The attributes of the PVT for the laptop and the wrist-worn device are shown in Table 1. In both devices, the test duration was 3 min and reaction time feedback was provided. The ISI denotes the period between the last response and the appearance of the next stimulus.

The rationale behind the decision to use the ISI of 2–10 s in the wrist-worn devices follows. The standard ISI for the original 10-min PVT is 2–10 s. Basner and colleagues shortened the duration of the test to 3 min and increased the signal rate by reducing the ISI to 1–4 s (Basner et al., 2011; Basner & Rubinstein, 2011). The authors noted that this increase in signal rate partially compensated for the reduction in the number of responses due to decreasing the trial duration from 10 to 3 min. However, this change of ISI, in conjunction with the trial duration, led to faster responses, an increased false start rate, and decreased lapse frequency in the brief PVT as compared to the original, 10-min one (Basner et al., 2011). These results suggest that short ISIs lead to PVT results that could be expected from more alert individuals. Furthermore, their results were obtained from participants in controlled

Table 1 Attributes of the PVT in the laptop and the wrist-worn device

PVT attributes	Computer-based PVT	Wrist-worn device-based PVT
Stimulus	Visual millisecond counter in red rectangular box	The appearance of the word “PUSH” in black digital letters presented in a low contrast LCD screen. When chosen, the red backlight will illuminate concurrently with the appearance of the word “PUSH.”
Test duration	3 min	3 min
ISI	Random: 1–4 s (both experiments)	Random: 2–10 s (1st experiment) Random: 1–4 s (2nd experiment)
Feedback	RT in milliseconds is displayed for 1 s (this period is part of the next ISI)	RT in milliseconds is displayed for 1 s.
False starts	“FS” is displayed for 1 s (this period is part of the next ISI)	“FS” is displayed for 1 s.
Time out without a response	After 30 s, “OVERRUN” is displayed for 1 s with an auditory signal to alert the participant. The response is counted as valid—that is, as a lapse with a response time of 30 s.	After 60 s, “DONE” is displayed for 1 s and the test session is closed. The response is not counted as valid.
Button fails to release	“BUTTON” is displayed after the response button has not been released for 3 s and a signal is continuously played back until the button is released. The new ISI starts once the button is released.	Device does not address this contingency. As soon as the button is pressed, the response is counted. Upon next stimulus, button is not recognized as depressed.
Wrong key pressed	“ERR” is displayed for 1 s if the wrong response key is pressed.	All available buttons are valid.

laboratory conditions. In contrast, crewmembers in our studies perform the wrist-worn PVT in their actual work environment, frequently with multiple individuals working in the same compartment, the existence of environmental noise, and so forth. These characteristics of the work environment may further affect the alertness level of the individuals performing the PVT. The other reason for choosing the ISI of 2–10 s is its operational validity. Even though the PVT is a simple task in comparison to the complexity of most operational tasks, a larger ISI denoting a less frequent task is more representative of the operational environment.

On the basis of these two concerns, the effect of the stimulus frequency on alertness and the external validity of the task, for the operational studies at NPS, we have used the wrist-worn device with an ISI of 2–10 s. The focus of this study was on the comparison of the wrist-worn 3-min PVT with relatively long ISIs (2–10 s) with the validated laptop-based 3-min PVT with shorter ISIs (1–4 s).

Procedures

Data were collected in two phases. In the first data collection, we focused on the main goal of this study—that is, comparing the wrist-worn PVT with an ISI of 2–10 s with the laptop PVT with an ISI of 1–4 s. We also assessed the effect of ambient light. The first experiment utilized a randomized, within-subjects repeated measures design with three factors. The first factor was the PVT device type (laptop – L, wrist-worn device – A). The second factor was the red backlight (BL) feature of

the wrist-worn device (BL = ON, BL = OFF). The third factor was ambient lighting, with two levels: a low-ambient-lighting condition similar to twilight (2–3 lux), and a normal office-lighting environment (300–400 lux). Ambient lighting was counterbalanced. Within each ambient-lighting condition, device order was also completely counterbalanced.

Participants first completed the study questionnaire. Participants were shown how to perform the PVT and were allowed one test trial with each device. To keep the reaction times as low as possible, participants were instructed to respond as soon as each stimulus appeared, but not to anticipate the target because that would yield a false start. Next, participants were randomly assigned to one of the twelve treatment groups of this experiment (Fig. 1). All participants performed six 3-min PVT trials, representing each of the conditions. Between trials, there was a 1-min break. In addition, a 5-min break was interposed between the ambient-lighting conditions. The length of time to complete the experiment was approximately 45 min for each participant. While performing the tests, participants were seated, wearing headphones to attenuate ambient noise. A researcher was present in the experimentation room to monitor the study.

On the basis of the findings of the first data collection, we then focused on the comparison of the wrist-worn and the laptop PVT in normal ambient-lighting conditions with both devices having an ISI between 1 and 4 s. Participants were randomly assigned to one of the six treatment groups. All participants performed three 3-min PVT trials, representing each of the conditions. Device order was completely counterbalanced.

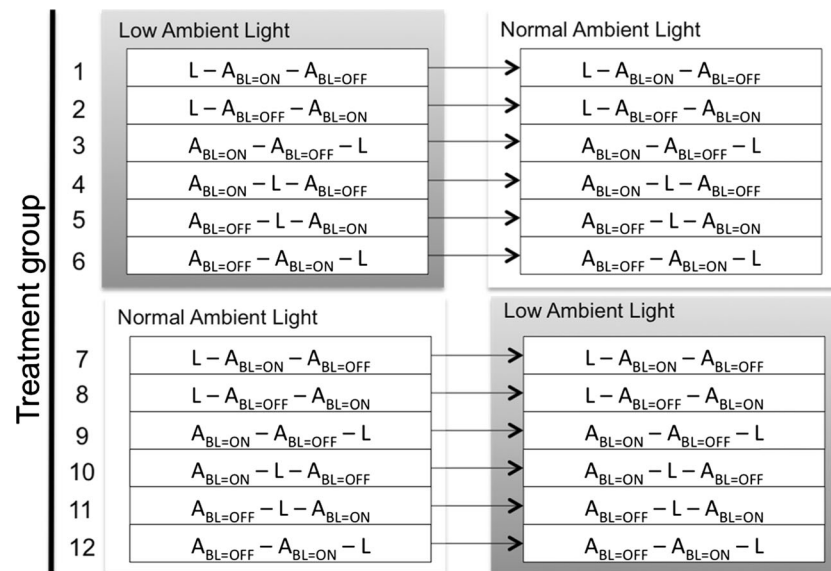


Fig. 1 Design and treatment groups in the first experiment

Statistical analysis

A PVT response was regarded as valid if the reaction time (RT) was greater than or equal to 100 ms and less than 30 s. Responses with RTs less than 100 ms were identified as false starts (errors of commission), and lapses were defined as RTs greater than or equal to either 355 or 500 ms (depending on the analysis). Using the PVT metrics proposed by Basner and Dinges (2011), our analysis included six PVT metrics: mean RT, mean response speed (i.e., the reciprocal reaction time, calculated as $1/RT \times 1,000$ and measured in $10^3 \cdot \text{ms}^{-1}$), fastest 10 % RT, slowest 10 % $1/RT$, percentage of 355-ms lapses combined with false starts, and percentage of 500-ms lapses combined with false starts. For all metrics, the response values were aggregated by trial. The validation criterion for each of the PVT metrics under focus was the absence of substantive differences between the wrist-worn device and the laptop.

First, we assessed the data for normality using the Shapiro–Wilk W test. With the exception of the reciprocal reaction times, the data were not normally distributed. Therefore, comparisons were based on nonparametric methods. Then, all variables underwent descriptive statistical analysis to identify anomalous entries and to calculate demographic characteristics. To assess the factors associated with PVT performance, we performed a mixed-effects analysis. The dependent variable was the reciprocal reaction time ($1/RT$) aggregated by trial. The fixed effects were PVT device type (wrist-worn device, laptop), ambient-lighting condition (low, normal), wrist-worn device backlight (on, off) nested within the device factor, and the sequence of ambient-lighting conditions. Subjects were included in the model as a random effect. Next, we compared the PVT metrics between conditions using matched-pairs Wilcoxon rank sum test (SAS Institute, 2007).

An alpha level of .05 was used to determine statistical significance. For multiple comparisons, post-hoc statistical significance was assessed using the Benjamini–Hochberg false discovery rate (BH-FDR) controlling procedure (Benjamini & Hochberg, 1995; Groppe, Urbach, & Kutas, 2011) at the .05 level. Statistical analysis was conducted with the JMP statistical software (JMP Pro 10; SAS Institute, Cary, NC). If not otherwise noted, the results in the text are presented as means (M) \pm standard deviations (SD).

Results

Participants ($n = 72$) reported sleeping 6.83 ± 0.95 h the night before the data collection. The average Epworth sleepiness score (ESS) score was 6.11 ± 4.09 . Neither reported sleep nor ESS scores differed between the two data collection groups (reported sleep: Wilcoxon rank sum test, $Z = 0.556$, $p = .578$; ESS score, Wilcoxon rank sum test, $Z = 0.322$, $p = .748$). Ten participants (five in each data collection phase) had an ESS score suggestive of elevated daytime sleepiness ($ESS > 10$). Given that the PVT metrics did not differ between the two ESS groups, all participants were combined and analyzed as a single group (Wilcoxon rank sum test, in all comparisons $p > .20$).

Main analysis

The cumulative distribution function (CDF) plots in Figs. 2 and 3 show the RTs of the PVT responses. Visual inspection of Figs. 2 and 3 shows two interesting patterns. In dim light conditions, the RTs of the wrist-worn device without backlight were considerably longer than those on the laptop and the

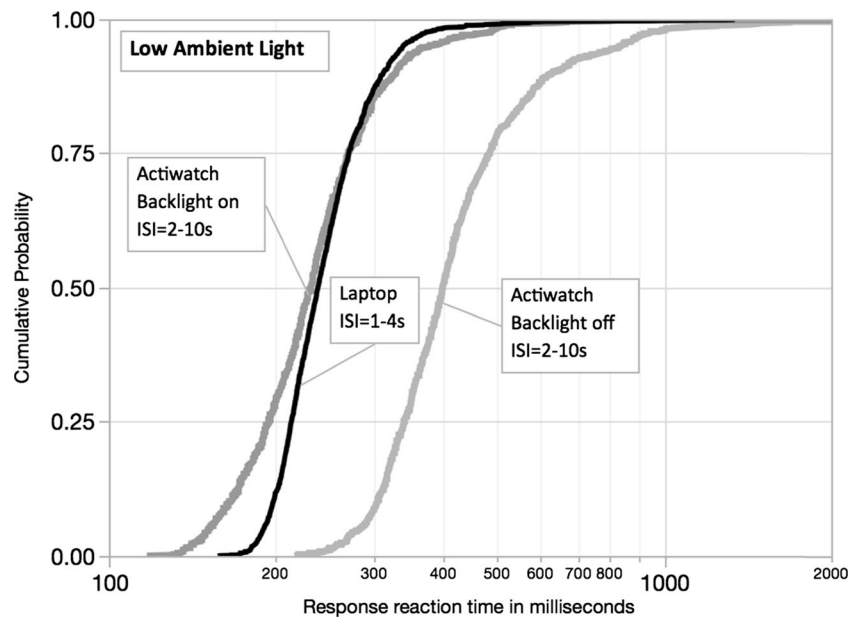


Fig. 2 Psychomotor vigilance test (PVT) responses in low ambient light. Wrist-worn PVT ISI = 2–10 s, laptop PVT ISI = 1–4 s

wrist-worn device with backlight. Second, when performing the PVT on the actiwatch with the backlight on, participants tended to have faster responses, as compared to their performance on both the laptop and the actiwatch with the backlight off. Specifically, in dim light, approximately 30 % of the responses on the wrist-worn device were between 100 and 200 ms, as compared to only 20 % of the responses on the laptop.

The remainder of our analysis focuses on PVT metrics aggregated by trial. Because of the shorter ISI (i.e., 1–4 s), the median number of responses per trial on the laptop was 48. In contrast, the longer ISI of 2–10 s for the wrist-worn device

led to a median of 21 responses per trial. Next, we performed a mixed-effects analysis to assess the factors associated with the reciprocal reaction time ($1/RT$). Visual inspection of the residual plots did not reveal any obvious deviations from homoscedasticity or normality. The results showed that the device type [$F(1, 176) = 64.6, p < .001$], the ambient-lighting condition [$F(1, 176) = 30.0, p < .001$], and the wrist-worn device backlight feature [$F(1, 176) = 292, p < .001$] all had main effects associated with the reciprocal RT. The corresponding multiple linear regression model explained 78 % of the overall variance. The reciprocal RTs were not associated, however, with the sequence of ambient-lighting conditions ($p > .70$).

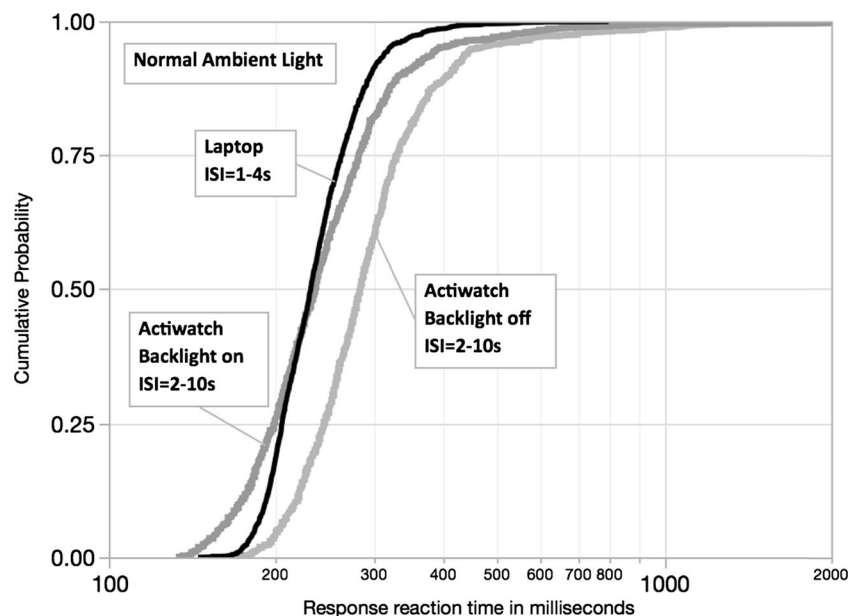


Fig. 3 PVT responses in normal ambient light. Wrist-worn PVT ISI = 2–10 s, laptop PVT ISI = 1–4 s

Table 2 PVT metrics by device-type, ambient-lighting, and backlight conditions

PVT metric	Ambient-lighting condition	Wrist-worn device with backlight off <i>M</i> ± <i>SD</i> (MD)	Wrist-worn device with backlight on <i>M</i> ± <i>SD</i> (MD)	Laptop <i>M</i> ± <i>SD</i> (MD)
Mean RT [ms]	Low	454 ± 144 (423)	250 ± 70.3 (246)	251 ± 27.3 (251)
	Normal	315 ± 75.0 (302)	267 ± 75.8 (253)	240 ± 27.2 (237)
Mean 1/RT [1,000/ms]	Low	2.47 ± 0.524 (2.47)	4.43 ± 0.919 (4.20)	4.16 ± 0.413 (4.07)
	Normal	3.52 ± 0.625 (3.52)	4.30 ± 0.840 (4.31)	4.32 ± 0.447 (4.31)
Fastest 10 % RT [ms]	Low	347 ± 89.0 (331)	194 ± 40.8 (193)	207 ± 19.1 (206)
	Normal	242 ± 38.1 (240)	195 ± 35.8 (187)	200 ± 19.5 (195)
Slowest 10 % 1/RT	Low	1.85 ± 0.557 (1.84)	3.41 ± 0.729 (3.42)	3.34 ± 0.436 (3.34)
	Normal	2.55 ± 0.786 (2.61)	3.19 ± 0.921 (3.14)	3.52 ± 0.472 (3.48)
Lapses 355 ms + False starts [%]	Low	69.7 ± 31.1 (82.6)	6.90 ± 16.1 (4.08)	5.42 ± 3.65 (4.17)
	Normal	19.3 ± 21.9 (13.3)	9.23 ± 12.8 (5.13)	5.0 ± 4.63 (4.08)
Lapses 500 ms + False starts [%]	Low	22.8 ± 25.6 (14.7)	2.11 ± 5.96 (0)	2.78 ± 2.18 (2.13)
	Normal	5.40 ± 8.32 (4.26)	3.86 ± 4.98 (0)	2.79 ± 2.57 (2.08)

Wrist-worn PVT ISI = 2–10 s, laptop PVT ISI = 1–4 s. *M* ± *SD* (MD): mean ± standard deviation (median)

We then assessed the PVT metrics by device, ambient-lighting condition, and wrist-worn device backlight (off, on). Table 2 shows these results.

Figure 4 further elaborates on the mean response speed by device type, ambient-lighting condition, and actiwatch backlight condition. Visual inspection of the data shows two patterns. In general, response speed is faster under normal lighting condition. Second, the median differences in response

speed between devices are more evident when comparing the actiwatch (backlight off) with the laptop or with the actiwatch with the backlight on.

From lower to upper, the horizontal lines in each box represent the 25th, 50th, and 75th percentiles, respectively. The upper end of the vertical line in each box extends to the outermost data point that falls within the distance computed as the 75th percentile + 1.5*interquartile range (IQR), whereas the

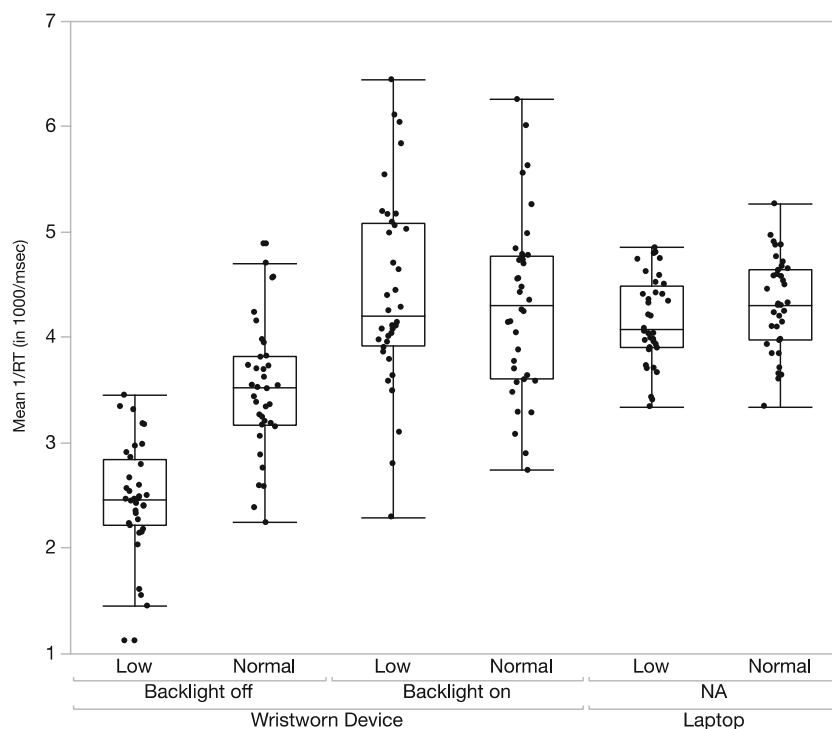


Fig. 4 PVT mean response speed, in 1,000/ms, by device type (wrist-worn or laptop), ambient-lighting condition (low or normal light), and actiwatch backlight condition (backlight off or on). Wrist-worn PVT ISI = 2–10 s, laptop PVT ISI = 1–4 s

lower end of the vertical line extends to the outermost data point that falls within the distance computed as the 25th percentile – 1.5*IQR. The IQR refers to difference between the 75th and 25th percentiles. If the data points did not reach the computed ranges, then the vertical lines were determined by the upper and lower data point values.

The patterns identified in Fig. 4 become more evident in the contrasts between all conditions (Table 3). Columns A to C include the median contrasts within each ambient-lighting condition, and show some interesting patterns. The median difference in RTs between the laptop and the wrist-worn device with the backlight on in low ambient-lighting conditions (column C) was less than 10 ms, which corresponds to less than 4 % of the RT found in the laptop PVT. Notably, no significant differences were identified in the percentages of 355-ms lapses combined with false starts. However, the absence of a backlight on the wrist-worn device had a considerable effect on the differences between the PVT metrics and those for the laptop (column B). Specifically, in low ambient-lighting conditions, the wrist-worn device shows an approximately 60 % increase in the mean and in the fastest 10 % of RTs, relative to the laptop. There is an approximately 76 % increase in the percentage of 355-ms lapses combined with false starts, whereas the increase is approximately 11 % in the percentage of 500-ms lapses combined with false starts. In the normal ambient-lighting condition, this pattern is evident but less pronounced. As compared to the laptop, the wrist-worn device shows an approximately 22 % increase in mean RTs, a 28 % increase in the fastest 10 % of RTs, and an 8 % increase in the percentage of 355-ms lapses combined with false starts.

No statistically significant difference was identified in the percentage of 500 ms lapses combined with false starts.

Columns D to F of Table 3 present the contrasts between ambient-lighting conditions. The pattern of results in column D shows that PVT performance in the wrist-worn devices is sensitive to ambient-lighting conditions when the backlight feature is off. In contrast, when the backlight is on (Column E of Table 3), PVT performance is not affected by ambient lighting. It is also notable that ambient-lighting conditions have a small but statistically significant effect on PVT performance when the laptop is used.

Elaborating on the effect of ISI

In the results of the first data collection, we determined that the wrist-worn PVT with an ISI of 2–10 s had statistically significant and substantive differences from the laptop PVT with an ISI of 1–4 s when the PVT was performed with the backlight off or when it was administered in low ambient-lighting conditions. On the basis of these results, we conducted a second experiment to identify the effect of ISI on these observed differences. Specifically, in the second data collection, the PVT was performed in normal ambient-lighting conditions with both devices having an ISI between 1 and 4 s.

The CDF plot in Fig. 5 shows the RTs of the PVT responses. The pattern of results is equivalent to the one observed in Fig. 3, in which the actiwatch and the laptop had different ISIs. That is, when performing the PVT on the actiwatch with the backlight on, participants tended to have faster responses, as compared to their performance on either

Table 3 Median contrasts in PVT metrics by device type, wrist-worn device backlight, and ambient-lighting condition

PVT metric	Ambient-lighting condition	Contrasts within each ambient-lighting condition			Contrasts between ambient-lighting conditions (low – normal)		
		Wrist-worn device with backlight off vs. Wrist-worn device with backlight on MD (95 % CI) [A]	Wrist-worn device with backlight off vs. Laptop MD (95 % CI) [B]	Wrist-worn device with backlight on vs. Laptop MD (95 % CI) [C]	Wrist-worn device with backlight off MD (95 % CI) [D]	Wrist-worn device with backlight on MD (95 % CI) [E]	Laptop MD (95 % CI) [F]
Mean RT [ms]	Low	170 (152, 198)***	162 (134, 203)***	-9.77 (-21.7, -0.36)*	111 (97.5, 124)***		10.8
	Normal	54.6 (41.3, 59.8)***	64.2 (49.4, 81.0)***				(6.19, 18.3)***
Mean 1/RT [1,000/ms]	Low	-1.99 (-2.08, -1.69)***	-1.63 (-1.83, -1.38)***	0.19 (0.02, 0.41)*	-1.03 (-1.19, -0.77)***		-0.15
	Normal	-0.84 (-0.93, -0.66)***	-0.84 (-0.99, -0.58)***				(-0.22, -0.11)***
Fastest 10 % RT [ms]	Low	135 (120, 158)***	123 (102, 149)***	-15.9 (-27.8, -6.82)**	95.7 (81.1, 111)***		7.51 (3.66, 9.88)***
	Normal	43.2 (37.8, 54.1)***	42.6 (31.3, 53.1)***				
Slowest 10 % 1/RT	Low	-1.58 (-1.72, -1.41)***	-1.40 (-1.62, -1.24)***	-0.20 (-0.54, -0.08)*	-0.71 (-0.94, -0.47)***		-0.16
	Normal	-0.66 (-0.85, -0.21)***	-0.98 (-1.02, -0.69)***				(-0.27, -0.06)**
Lapses 355 ms + False starts [%]	Low	72.6 (54.9, 81.4)***	76.3 (61.1, 86.8)***		55.3 (35.0, 70.4)***		-3.92 (-4.76, 0)**
	Normal	5.74 (4.35, 9.05)***	8.01 (2.96, 17.0)***				
Lapses 500 ms + False starts [%]	Low	13.5 (9.09, 20.0)***	10.9 (5.36, 20.0)***	-2.08 (-2.13, 0)**	9.52 (4.55, 18.0)***		0 (-4.35, 0)*
	Normal						

Wrist-worn PVT ISI = 2–10 s, laptop PVT ISI = 1–4 s. MD (95 % CI) = Median (95 % confidence interval of the median). Inclusion criterion: All shown contrasts are statistical significant according to post-hoc analysis with the BH-FDR controlling procedure. Empty cells denote results that are statistically not significant. Statistical significance of the shown contrasts, based on the Wilcoxon signed rank test: * $p < .05$, ** $p < .01$, *** $p < .001$

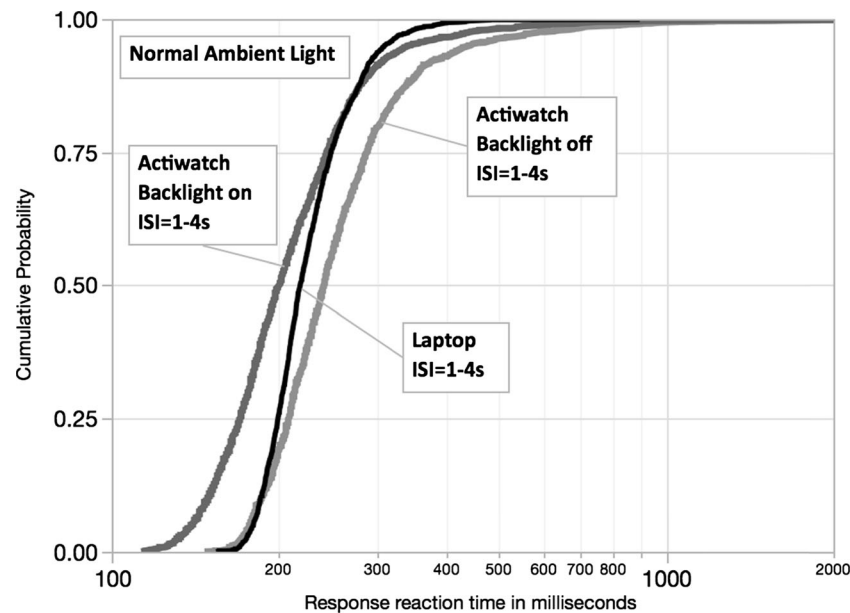


Fig. 5 PVT responses in normal ambient light. Both devices with ISI = 1–4 s

the laptop or the actiwatch with the backlight off. It is notable, however, that responses on the wrist-worn device between 100 and 200 ms increased from approximately 30 % in the ISI = 2–10-s condition to approximately 50 % when the ISI was 1–4 s.

Next, we performed a mixed-effects analysis over the entire data set to assess whether ISI was associated with the reciprocal reaction time ($1/RT$). The fixed effects were ISI, device type (wrist-worn device, laptop), ambient-lighting condition (low, normal), and wrist-worn device backlight (on, off) nested within the Device Type factor. Subjects were included in the model as a random effect. Adjusted for the other statistically significant factors (device type, ambient light, and backlight feature), ISI was a significant predictor of the reciprocal reaction time [$F(1, 317.6) = 39.7, p < .001$], with a shorter ISI leading to faster response speeds. Table 4 shows the PVT metrics by device and wrist-worn device backlight (off, on). For both devices, the ambient-lighting conditions were normal and the ISI was between 1 and 4 s.

Discussion

In this study, we focused on the differences in the 3-min PVT performance metrics when it is performed on a wrist-worn device with an ISI of 2–10 s versus the validated laptop-based 3-min PVT with an ISI of 1–4 s. Our results show that, when the wrist-worn device backlight is on, the median difference in reaction times between the laptop and the wrist-worn device in low ambient-lighting conditions was less than 10 ms—which corresponds to less than 4 % of the RT found in the laptop PVT. Furthermore, the corresponding median differences in response speed and in the percentage of either 355-ms or 500-ms lapses combined with false starts are less than 4.5 %. The 10-ms median difference in RTs is not considered operationally significant (Khitrov et al., 2014). In general, these small differences are statistically significant only in the low ambient-lighting conditions. Only the difference in the response speed of the slowest 10 % of the responses remains statistically significant in the normal light conditions.

Table 4 PVT metrics by device type and backlight on/off

PVT metric	Wrist-worn device with backlight off $M \pm SD$ (MD)	Wrist-worn device with backlight on $M \pm SD$ (MD)	Laptop $M \pm SD$ (MD)
Mean RT [ms]	272 \pm 55.1 (258)	225 \pm 49.3 (216)	230 \pm 24.2 (226)
Mean $1/RT$ [1,000/ms]	4.08 \pm 0.61 (4.14)	5.00 \pm 0.81 (5.00)	4.49 \pm 0.412 (4.51)
Fastest 10 % RT [ms]	204 \pm 30.4 (197)	164 \pm 28.3 (158)	194 \pm 15.5 (193)
Slowest 10 % $1/RT$	2.95 \pm 0.62 (2.95)	3.62 \pm 0.70 (3.09)	3.64 \pm 0.448 (3.58)
Lapses 355 ms + False starts [%]	11.0 \pm 12.7 (11.0)	5.17 \pm 7.53 (4.12)	4.42 \pm 3.87 (4.09)
Lapses 500 ms + False starts [%]	4.14 \pm 5.08 (4.14)	2.37 \pm 3.05 (2.11)	2.80 \pm 2.52 (2.08)

For both devices, ISI = 1–4 s. $M \pm SD$ (MD): mean \pm standard deviation (median)

However, the standard deviation for most PVT metrics for the wrist-worn devices was considerably larger when compared to data obtained from the laptop version. For RTs, the standard deviation was between 50 ms (ISI = 1–4 s) and 72 ms (ISI = 2–10s), as compared to approximately 26 ms on the laptop, or 29 ms of intraparticipant variability in the original PVT-192 results from sleep-satiated participants (as reported by Khitrov et al., 2014, based on their analysis of the baseline sessions of Rupp, Wesensten, & Balkin, 2012). Wrist-worn device variability was increased in both ambient-lighting and ISI conditions; we believe that this finding can be attributed predominantly to differences in the mechanical characteristics of the buttons used to respond in the two devices.

Overall, when using the wrist-worn device with the backlight on, our results suggest that the average PVT performance does not seem to be sensitive to changes in ambient-lighting conditions. This finding can be explained if we consider the stimulus characteristics. In the wrist-worn PVT, the stimulus is the word “PUSH” in black digital letters, which is presented on a low-contrast LCD screen. Depending on the intensity of the ambient light, therefore, performing the PVT on the wrist-worn device may include a signal detection component. We postulate that the latter is driving PVT performance when ambient lighting is low and the backlight is off. The use of the backlight feature, however, introduces a second visual cue enhancing the stimulus. The reason is that the backlight turns on concurrently with the “PUSH” stimulus. Hence, participants tend to respond to the backlight instead of the letters. This explanation is supported by the higher percentage of fast responses (100–200 ms) in the wrist-worn device with the backlight on than when using either the same device with the backlight off or the laptop. This pattern is consistent in both ISI conditions—that is, when the wrist-worn and the laptop have either the same or different ISIs.

On the basis of our results, when collecting data in the field with the AMI device, we recommend turning the backlight on and using an ISI of 2–10 s for the 3-min PVT. The use of the backlight may alleviate the problem of inconsistent ambient-lighting conditions in the field, and hence, produce less variable PVT results that are more comparable to those of the laptop PVT. The PVT embedded in a wrist-worn device also provides a simple, less disruptive, and rapid operational test of psychomotor vigilance performance for occupations outside the military, such as security-related occupations, first responders and emergency medical teams, power grid and plant operators doing shift work, and air traffic controllers.

Limitations

Several limitations should be taken into account when interpreting our results. This study assessed the differences in PVT performance between devices in a relatively young and healthy population (90 % were less than 45 years old).

Participants also estimated the amount of sleep they received on the previous night. Follow-up studies should use objective measures to assess participants' sleep. We did not use an additional calibration device to assess the accuracy of the timed reactions (Khitrov et al., 2014); different devices with different screen characteristics and stimulus presentation may yield different results.

One of the criticisms of the PVT is the high presentation rate of stimuli. In general, vigilance tasks in the operational environment are characterized by their infrequent, or even rare, occurrence, which would be represented better by longer ISIs (Wolfe, Horowitz, & Kenner, 2005; Van Wert, Horowitz, & Wolfe, 2009). In contrast, much of the PVT literature has focused on ISIs 1–10 s in length. Future studies should compare the differences in PVT metrics when both the wrist-worn device and the laptop have the same but longer ISIs—for example, 10 to 30 s.

Author note This work was funded by and prepared for the Advanced Medical Development Program, Naval Medical Research Center, 503 Robert Grant Avenue, Silver Spring, Maryland 20910. The authors also acknowledge the effort and valuable comments of the reviewers of the manuscript. Disclosure statement: This work is not an industry-supported study and did not use off-label or investigational items. The authors have no conflicts of interest.

References

- Balkin, T., Thorne, D., Sing, H., Thomas, M., Redmond, D. P., Wesensten, N. J., . . . Belenky, G. (2000). *Effects of sleep schedules on commercial motor vehicle driver performance* (Final, Report No. DOT-MC-00-133). Springfield, VA: U.S. Department of Transportation.
- Basner, M., & Dinges, D. F. (2011). Maximizing sensitivity of the Psychomotor Vigilance Test (PVT) to sleep loss. *Sleep, 34*, 581–591.
- Basner, M., Mollicone, D., & Dinges, D. F. (2011). Validity and sensitivity of a brief psychomotor vigilance test (PVT-B) to total and partial sleep deprivation. *Acta Astronautica, 69*, 949–959.
- Basner, M., & Rubinstein, J. (2011). Fitness for duty: A 3-minute version of the Psychomotor Vigilance Test predicts fatigue-related declines in luggage-screening performance. *Journal of Occupational and Environmental Medicine, 53*, 1146–1154. doi:10.1097/JOM.0b013e31822b8356
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B, 57*, 289–300.
- Dinges, D. F., Pack, F., Williams, K., Gillen, K. A., Powell, J. W., Ott, G. E., . . . Pack, A. I. (1997). Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4–5 hours per night. *Sleep, 20*, 267–277.
- Dinges, D. F., & Powell, J. W. (1985). Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers, 17*, 652–655. doi:10.3758/BF03200977
- Doran, S. M., Van Dongen, H. P. A., & Dinges, D. F. (2001). Sustained attention performance during sleep deprivation: Evidence of state instability. *Archives Italiennes de Biologie, 139*, 253–267.

- Durmer, J. S., & Dinges, D. F. (2005). Neurocognitive consequences of sleep deprivation. *Seminars in Neurology*, *25*, 117–129.
- Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields II: Simulation studies. *Psychophysiology*, *48*, 1726–1737. doi:10.1111/j.1469-8986.2011.01272.x
- Jewett, M. E., Dijk, D. J., Kronauer, R. E., & Dinges, D. F. (1999). Dose-response relationship between sleep duration and human psychomotor vigilance and subjective alertness. *Sleep*, *22*, 171–179.
- Khitrov, M. Y., Laxminarayan, S., Thorsley, D., Ramakrishnan, S., Rajaraman, S., Wesensten, N. J., & Reifman, J. (2014). PC-PVT: A platform for psychomotor vigilance task testing, analysis, and prediction. *Behavior Research Methods*, *46*, 140–147. doi:10.3758/s13428-013-0339-9
- Kribbs, N. B., & Dinges, D. F. (1994). Vigilance decrement and sleepiness. In J. R. Harsh & R. D. Ogilvie (Eds.), *Sleep onset mechanisms* (pp. 113–125). Washington, DC: American Psychological Association.
- Lamond, N., Dawson, D., & Roach, G. (2005). Fatigue assessment in the field: Validation of a hand-held electronic psychomotor vigilance task. *Aviation, Space, and Environmental Medicine*, *76*, 486–489.
- Lamond, N., Jay, S. M., Dorrian, J., Ferguson, S. A., Roach, G. D., & Dawson, D. (2008). The sensitivity of a palm-based psychomotor vigilance task to severe sleep loss. *Behavior Research Methods*, *40*, 347–352. doi:10.3758/BRM.40.1.347
- Miller, N. L., Matsangas, P., & Kenney, A. (2012). The role of sleep in the military: Implications for training and operational effectiveness. In J. H. Laurence & M. D. Matthews (Eds.), *The Oxford handbook of military psychology* (pp. 262–281). New York, NY: Oxford University Press.
- Rosekind, M. R., Graeber, R. C., Dinges, D. F., Connel, L. J., Rountree, M. S., & Spinweber, C. L. (1994). *Crew factors in flight operations IX: Effects of planned cockpit rest on crew performance and alertness in long-haul operations* (NASA TM-108839). Moffett Field, CA: NASA Ames Research Center.
- Rosekind, M. R., Gregory, K. B., & Mallis, M. M. (2006). Alertness management in aviation operations: Enhancing performance and sleep. *Aviation, Space and Environmental Medicine*, *77*, 1256–1265.
- Rupp, T. L., Wesensten, N. J., & Balkin, T. J. (2012). Trait-like vulnerability to total and partial sleep loss. *Sleep*, *35*, 1163–1172. doi:10.5665/sleep.2010
- SAS Institute. (2007). *JMP statistics and graphics guide, release 7*. Cary, NC: SAS Institute.
- Shattuck, N. L., Matsangas, P., & Brown, S. (2015). *A comparison between the 3/9 and the 5/10 watchbills* (Technical Report Report No. NPS-OR-15-006). Monterey, CA: Naval Postgraduate School.
- Shattuck, N. L., Matsangas, P., & Powley, E. H. (2015). *Sleep patterns, mood, psychomotor vigilance performance, and command resilience of watchstanders on the “five and dime” watchbill* (Technical Report No. NPS-OR-15-003). Monterey, CA: Naval Postgraduate School.
- Shattuck, N. L., Matsangas, P., & Waggoner, L. (2014, October). *Assessment of a novel watchstanding schedule on an operational US Navy vessel*. Paper presented at the Human Factors and Ergonomics Society (HFES) 58th Annual Meeting, Chicago, IL.
- Shattuck, N. L., Waggoner, L. B., Young, R. L., Smith, C. S., & Matsangas, P. (2014). Shiftwork practices in the United States Navy: A study of sleep and performance in watchstanders aboard the USS Jason Dunham. *Sleep*, *37*(Abstract Suppl.), A78.
- Thorne, D., Johnson, D. E., Redmond, D. P., Sing, H. C., Belenky, G., & Shapiro, J. M. (2005). The Walter Reed palm-held psychomotor vigilance test. *Behavior Research Methods*, *37*, 111–118. doi:10.3758/BF03206404
- Van Wert, M. J., Horowitz, T. S., & Wolfe, J. M. (2009). Even in correctable search, some types of rare targets are frequently missed. *Attention, Perception, & Psychophysics*, *71*, 541–553. doi:10.3758/APP.71.3.541
- Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005). Cognitive psychology: Rare items often missed in visual searches. *Nature*, *435*, 439–440. doi:10.1038/435439a
- Wyatt, J. K., Dijk, D. J., Ronda, J. M., Jewett, M. E., Powell, J. W., Dinges, D. F., & Czeisler, C. A. (1997). Interaction of circadian- and sleep/wake homeostatic-processes modulate psychomotor vigilance test (PVT) performance [Abstract]. *Sleep Research*, *26*, 759.