Check for updates

Agreement between the 3-minute Psychomotor Vigilance Task (PVT) Embedded in a Wrist-worn Device and the Laptop-based PVT

Panagiotis Matsangas and Nita Lewis Shattuck
Operations Research Department, Naval Postgraduate School, Monterey, CA

The study assesses the agreement between the 3-minute version of the Psychomotor Vigilance Task (PVT) with an interstimulus interval (ISI) of 2 to 10 seconds and the validated 3-minute laptop-based PVT (ISI=1-4 seconds). The experiment utilized a randomized, within-subject, repeated-measures design with three factors (PVT device type, the backlight feature of the wrist-worn device, ambient lighting). Results show the differences in reaction times (RT) between devices are incrementally associated with the magnitude of the RTs. These differences tend to be in opposing directions when the backlight feature in the wrist-worn device is on. That is, RTs in the wrist-worn device tend to be faster compared to the laptop for (on average) faster individuals, whereas (on average) slower individuals tend to do better in the laptop compared to the wrist-worn device. The proportional bias introduced by the wrist-worn device compared to the laptop makes it difficult to translate individual RTs between different devices. The proportional bias, however, may work in favor for detecting differences between slow and fast RTs.

INTRODUCTION

Background

The Psychomotor Vigilance Test (PVT) is a widely used reaction time tool to assess behavioral alertness and vigilance, i.e., the ability to pay attention for prolonged periods of time to detect infrequently occurring stimuli. The original version of the PVT assumes that the test is performed using a computer and the task duration is 10 minutes in length (Dinges & Powell, 1985).

Using the 10-minute computer version of the PVT, however, is not practical in field studies on naval vessels (Matsangas, Shattuck, & Brown, In press). Shorter versions of the PVT, such as the 3-minute version, have been shown to reliably detect the effects of sleep loss (Basner & Rubinstein, 2011). In field studies conducted by faculty and students from the Naval Postgraduate School, we use the 3-minute version of the PVT embedded in a actigraph, i.e., a wrist-worn device to assess sleep patterns (Miller, Matsangas, & Kenney, 2012; Shattuck, Matsangas, & Dahlman, 2018; Shattuck, Matsangas, Mysliwiec, & Creamer, In press). The extensive use of this configuration has produced positive results. Compliance for taking the PVT according to the study protocol has almost tripled compared to earlier studies using the laptop-based PVT. For example, in a recent study collecting PVT data from crewmembers on four ships while underway, approximately 60% of the participating crewmembers provided PVT data useful for analysis. Data from our field studies were reasonable and results showed that the configuration we used could distinguish psychomotor vigilance performance differences between watch standing schedules used on U.S. Navy ships (Shattuck, Matsangas, & Brown, 2015; Shattuck, Matsangas, & Powley, 2015; Shattuck, Matsangas, & Waggoner, 2014).

In field studies, however, the PVT is used in conditions deviating from the controlled laboratory conditions in which the test has been validated. For example, one issue of concern was the ambient light conditions when crewmembers perform the PVT. The short 1 to 4 seconds inter-stimulus interval (ISI, i.e., the period between the last response and the appearance of the next stimulus) used in the validated laptop-based PVT was also of concern.

Compared to the standard ISI=2-10 seconds in the 10-minute PVT, Basner and colleagues reduced the ISI 1 to 4 seconds in the 3-minute PVT to increase signal rate (Basner, Mollicone, & Dinges,

Not subject to U.S. copyright restrictions. DOI 10.1177/1541931218621151

2011; Basner & Rubinstein, 2011). This change led to PVT results expected from more alert individuals, e.g., faster responses (Basner et al., 2011). In contrast, the wrist-worn 3-minute PVT we were using had ISI=2-10 seconds. For a detailed discussion regarding the rationale to use the ISI=2-10 seconds in the wrist-worn devices refer to Matsangas et al. (In press).

With these concerns in mind, we conducted an initial validation study of the wrist-worn PVT (Matsangas et al., In press). Results showed that, regardless of the ambient light conditions, the overall median difference in PVT metrics between the laptop and the wrist-worn PVT was less than 4.5% when the backlight was on. PVT metrics, however, exhibited increased variability in the wrist-worn device compared to the laptop.

This paper continues the assessment of the wrist-worn PVT by focusing on the variability of the differences between the wrist-worn and validated laptop-based PVT. Specifically, the focus of this study is to assess the differences in reaction time between the 3-minute PVT embedded in a wrist-worn device (ISI= 2-10 seconds) and the validated laptop-based PVT (ISI= 1-4 seconds) using the Bland-Altman method (Altman & Bland, 1983; Bland & Altman, 1986).

This work was supported by the Naval Medical Research Center's Advanced Medical Development Program.

METHOD

Participants

Thirty-six healthy individuals (77.8% males), on average 36.0±7.47 years of age, from the Naval Postgraduate School (NPS) volunteered to participate to the study. The study protocol was approved by the Institutional Review Board at the Naval Postgraduate School.

Equipment and Instruments

Psychomotor vigilance performance data were collected using two devices: a) a wrist-worn device (Motionlogger Watch) with an embedded version of the PVT (Ambulatory Monitoring, Inc., Ardsley, NY), and b) a validated laptop-based PVT (PULSAR Informatics, Philadelphia, PA).

In both devices, the PVT trial duration was 3 minutes and feedback of the reaction time was provided on the device's screen for 1 second after the response. Denoting the period between the last response and the appearance of the next stimulus, the inter-stimulus interval (ISI) was set to 2-10 seconds in the wrist-worn device and 1-4 seconds in the laptop.

Procedures

The experiment utilized a randomized, within-subject, repeated-measures design with three factors (PVT device type, the backlight feature of the wrist-worn device, ambient lighting). Ambient lighting had two levels: a low ambient lighting condition similar to twilight (2 – 3 lux), and a normal office lighting environment (300 – 400 lux). Each participant performed six 3-minute PVT trials, three trials in normal light and three trials in low ambient light. Ambient lighting was counterbalanced. Within each ambient light condition, device order was also completely counterbalanced. Detailed information is presented elsewhere (Matsangas et al., In press).

Analysis

Analysis was based on the Bland–Altman method (Altman & Bland, 1983; Bland & Altman, 1986, 1999) for comparing different measurement methods. Preliminary analysis showed that the mean difference between devices was associated with the magnitude of the measurements in all four conditions of the experiment (low/normal ambient light and backlight on or off). Therefore, we used the regression approach for non-uniform differences and calculated the 95% limits of agreement (Bland & Altman, 1999).

Statistical analysis was conducted with JMP statistical software (JMP Pro 13; SAS Institute; Cary, NC). A PVT response was regarded as valid if the reaction time (RT) was greater than or equal to 100 milliseconds (ms) and less than 30 seconds.

RESULTS

Table 1 shows the difference in reaction time between the wrist-worn PVT and the laptop PVT. When the backlight feature is on, the reaction time differences between the two devices are on average small (Matsangas et al., In press).

Table 1. Reaction time differences between devices.

Backlight feature (wrist-worn device)	Ambient light	Median difference [A, B]
Off	Low	162 (47.5%)
	Normal	64.2 (24.8%)
On	Low	-9.77 (-4.09%)
	Normal	8.78 (3.41%)

[A] Difference in RTs = Wrist-worn device RT minus laptop RT
[B] Median difference in RT presented in milliseconds followed by the percentagewise difference in parenthesis

Figures 1 and 2 show the scatter diagrams of the RTs in the two devices. The solid black lines indicate the equality (45 degrees) lines. It is evident that when the backlight is off, there is a systematic bias in the wrist-worn RTs (responses fall well above the equality line). In contrast, when the backlight is on, PVT responses are better clustered around the equality line.

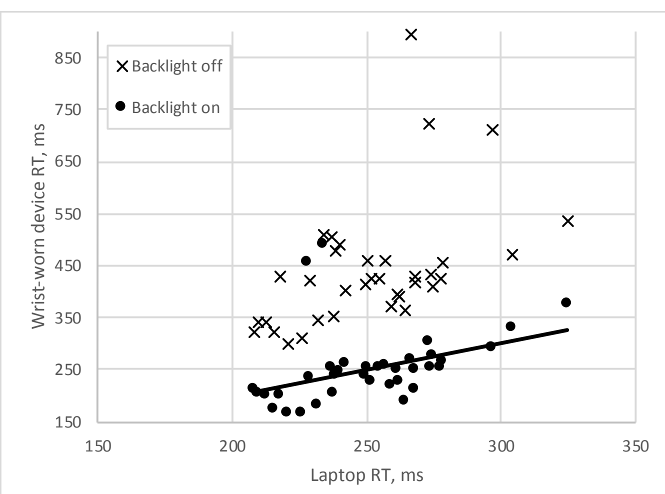

Figure 1. Reaction time in low ambient light

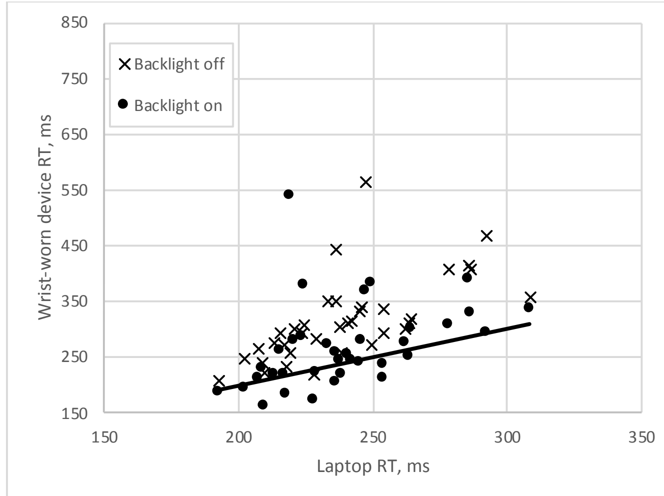

Figure 2. Reaction time in normal ambient light

Figures 3 to 6 show that Bland-Altman plots of the RTs in the four experimental conditions (low/normal ambient light, backlight on/off). The regression line is denoted by the solid black line. The dotted lines represent the regression-based 95% limits of agreement. It is evident that in all conditions, there a proportional bias, i.e., there is an incremental relationship between the RT differences and the magnitude of the RTs. These differences tend to be in opposing directions when the backlight is on. Specifically, negative differences (i.e., faster responses in the wrist-worn device compared to the laptop) are associated with short RTs, whereas positive differences (i.e., slower responses in the wrist-worn device compared to the laptop) are associated with long RTs. Each figure has different range in the the x and y axes for better depiction of differences.

DISCUSSION

Results showed that the 3-minute PVT embedded in the wrist-worn device with ISI=2-10 seconds is associated with a proportional bias compared to the validated 3-minute, laptop-based, PVT with ISI=1-4 seconds. When the backlight is off, this incremental relationship between the RT differences and the magnitude of the RTs is in the same direction, i.e., the wrist-worn PVT has consistently longer RTs compared to the laptop.

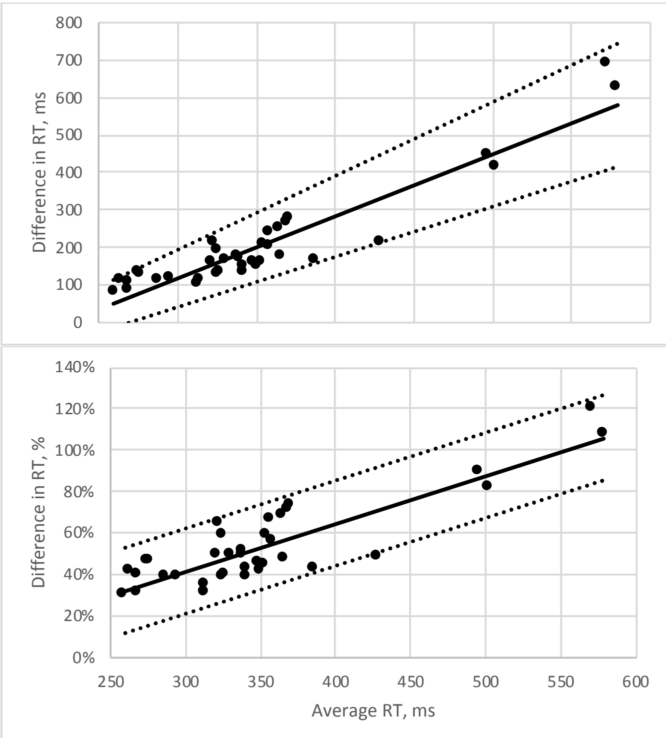

Figure 3. Absolute and percentagewise differences in PVT reaction times in low ambient light conditions. Actiwatch with backlight off.

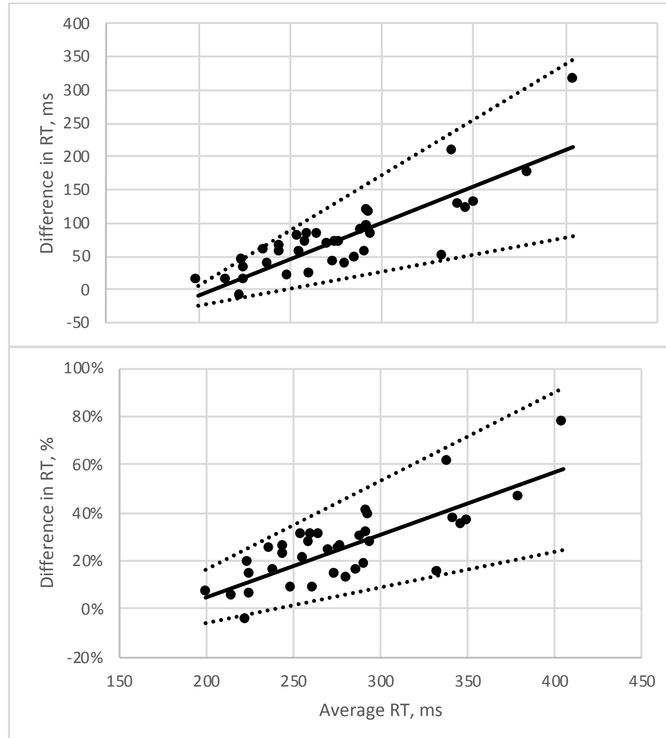

Figure 5. Absolute and percentagewise differences in PVT reaction times in normal ambient light conditions. Actiwatch with backlight off.

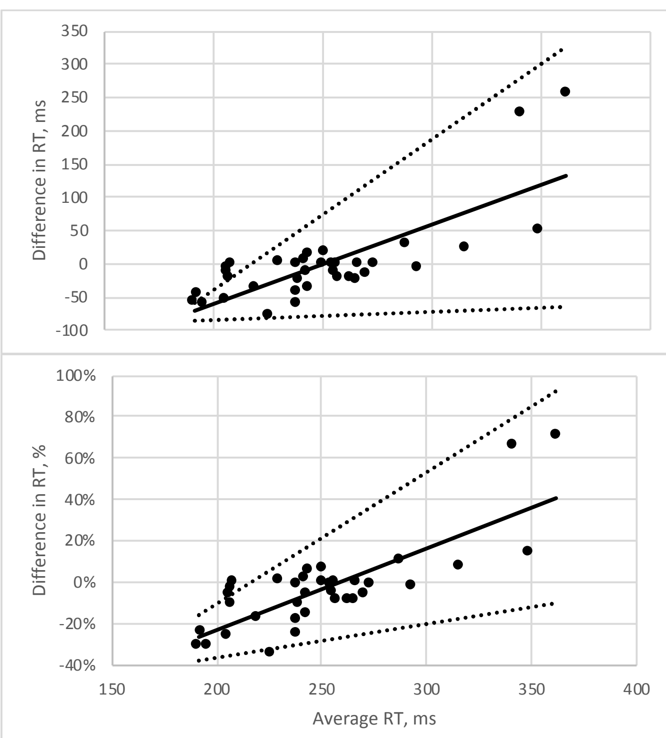

Figure 4. Absolute and percentagewise differences in PVT reaction times in low ambient light conditions. Actiwatch with backlight on.

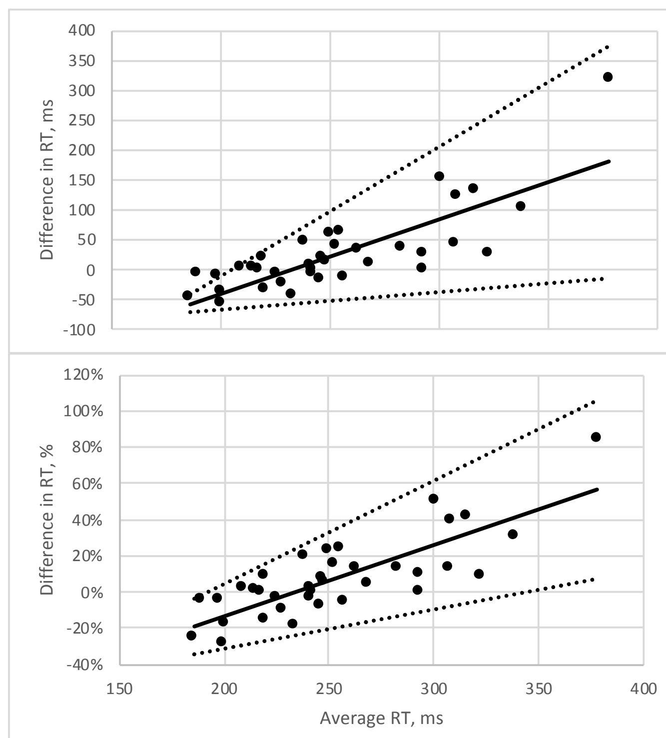

Figure 6. Absolute and percentagewise differences in PVT reaction times in normal ambient light conditions. Actiwatch with backlight on.

When the backlight is on, wrist-worn RTs are, in general, faster compared to the corresponding RTs when the backlight is off. This finding results from having differences in RTs between devices in opposing directions. That is, RTs in the wrist-worn device tend to be faster compared to the laptop for faster individuals who are, on average, faster, whereas individuals who are, on average, slower tend to do better on the laptop compared to their performance on the wrist-worn device.

We postulate that our findings can be attributed to two factors. First, the ISI in the wrist-worn device is longer compared to the ISI of the laptop. All other factors being equal, longer ISIs are associated with longer RTs (Basner et al., 2011). Second, there is a backlight in the wrist-worn device. When this feature is on, the screen lights when the word "PUSH" (the primary stimulus) is presented. The backlight not only improves the contrast of the screen, but also acts as a second additional visual stimulus.

Combined with earlier analysis, findings from this study support the use of the 3-minute PVT embedded in the wrist-worn device. Even though the proportional bias of the device we identified makes it difficult to translate individual RTs between different devices, such comparisons for large samples may be of value. Furthermore, the proportional bias when the backlight is on may work in favor for detection of differences between slow and fast RTs.

Study limitations

The sample of 36 participants may be small for the Bland-Altman analysis, leading to wide limits of agreements. To avoid fatigue effects, participants performed only one PVT trial in each experimental condition. Future efforts should include a larger sample of participants with multiple replications in each experimental condition (Bland & Altman, 1999).

Lastly, the study participants were working typical office hours and were neither shiftworkers nor severely sleep-deprived. Even though a few reported elevated daytime sleepiness, their PVT results did not differ substantially from the rest of participants. In short, the study participants were considered in general well rested. Follow-up studies should use participants with fatigue levels comparable to those encountered in actual naval operational conditions.

Acknowledgements

This work was funded by, and prepared for, the Advanced Medical Development Program; Naval Medical Research Center; 503 Robert Grant Avenue, Silver Spring, MD 20910.

LIST OF REFERENCES

Altman, D. G., & Bland, J. M. (1983). Measurement in medicine: The analysis of method comparison studies. *Statistician*, *32*(3), 307-317.

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*(11-12), 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*(10), 1146-1154.

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, *1*(8476), 307-310.

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, *8*(2), 1350-1160.

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*(6), 652–655.

Matsangas, P., Shattuck, N. L., & Brown, S. (In press). Preliminary validation study of the 3-minute wrist-worn Psychomotor Vigilance Test. *Behavior Research Methods*.

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: Oxford University Press.

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., & Dahlman, A. S. (2018). Sleep and fatigue issues in military operations. In E. Vermetten, A. Germain, & T. Neylan (Eds.), *Sleep and Combat related PTSD*: Springer.

Shattuck, N. L., Matsangas, P., Mysliwiec, V., & Creamer, J. L. (In press). The role of sleep in human performance and well-being. In D. Schnyer & M. D. Matthews (Eds.), *The Cognitive and Behavioral Neuroscience of Human Performance in Extreme Settings*. New York: Oxford University Press.

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 27-31). *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.