Published in final edited form as:

Sleep Health. 2018 October ; 4(5): 387–389. doi:10.1016/j.sleh.2018.08.003.

Addressing the Need for Validation of a Touchscreen Psychomotor Vigilance Task: Important Considerations for Sleep Health Research

Michael A. Grandner, PhD MTR¹, Nathaniel F. Watson, MD, MSc², Matthew Kay, PhD³, Demi Ocano¹, and Julie A. Kientz, PhD⁴

¹Sleep and Health Research Program, Department of Psychiatry, University of Arizona College of Medicine, Tucson, AZ

²Department of Neurology and UW Medicine Sleep Center, University of Washington School of Medicine, Seattle, WA

³School of Information, University of Michigan

⁴Department of Human Centered Design and Engineering, University of Washington, Seattle, WA

The psychomotor vigilance task (PVT) ¹ is a validated measure of sustained attention and vigilance and the gold standard for detecting objective neurobehavioral impairment due to sleep loss². The PVT has been used to assess impairment of vigilance and sustained attention from total and partial sleep deprivation^{3, 4} and circadian dysregulation⁵, and has also been used as a method for estimating daytime sleepiness ⁶. The standard PVT duration is 10 minutes in duration, although shorter versions have been validated for specific applications/platforms ^{7, 8}.

The standard laboratory-based PVT device (the PVT-192, marketed by Ambulatory Monitoring, Inc., Ardsley, NY) has several practical limitations. The test equipment can be expensive and includes outdated legacy electronic hardware. The PVT-192 equipment is not easily portable, which limits its utility in field research. To date, there have been several attempts to address portability. The Walter Reed Palm PVT, a software program for the legacy Palm Pilot device, successfully demonstrated the proof-of-concept of a more portable PVT test ⁹. However, Palm Pilots are no longer manufactured and the Palm PVT software is no longer supported. More recently, the PVT has been translated into a touchscreen interface by multiple groups¹⁰⁻¹⁴, though validation studies are limited. Also, these applications are generally not designed to be easily disseminated in large research projects.

In addition to portability, implementation of a PVT in community health research needs to address a few operational issues. First, transitioning from analog buttons to digital touch

CORRESPONDING AUTHOR: Dr. Michael Grandner, University of Arizona Sleep and Health Research Program, 1501 N Campbell Ave, PO Box 245002, Tucson, AZ 95724-5002; grandner@email.arizona.edu; (520) 626-6346.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

screens affects PVT implementation. In the traditional PVT approach, an individual rests their finger on the button, pressing as soon as they see the stimulus producing an immediate reaction time. However, touchscreen functionality is such that the individual cannot rest their hand on the screen without triggering a response. Prior work by our group examined several approaches that could capture reaction time on a touchscreen, including goal-crossing, finger lift, finger tilt, touch down, and home button press¹⁴. We found users strongly preferred screen touch but this created a problem – the added time it takes a finger to traverse space and press the screen (and the lack of standardization of that distance) creates a systematic slowing in reaction time values¹⁴. Any touchscreen solution and analysis would need to account for this.

Second, various devices present hardware and software permissions issues that disallow unfettered access to a reliable timekeeping clock or the ability to ensure minimal delay between the software's stimulus presentation and its display on screen. Further, these devices lack a reliable rate of polling touchscreen sensors (i.e., when a finger touches the screen, how quickly will the system register that contact), and some of these devices do not allow programs to access the screen refresh rate (how quickly an instruction to display something on the screen is actually displayed). All of these concerns lead to challenges in implementing a valid touchscreen PVT application. It is unclear (and likely variable across devices) the degree to which problems with touchscreen implementation of the PVT are related to systematically slower response times due to the finger traversing space vs touchscreen timing issues at the device level.

To address these concerns, we developed the PVT-Touch (<http://pvttouch.com>) for the Android operating system, the most common operating system among smartphones in the US¹⁵. To validate the PVT-Touch relative to the gold-standard PVT-192, participants (N=60 adults age 25-60) were recruited through flyers and online advertisements for a study of health outcomes associated with habitual sleep duration. Subjects completed extensive screening questionnaires and were free from major comorbidities that impact sleep including depression or other psychiatric illness, insomnia disorder, or sleep apnea. This study was approved by the IRB of both the University of Pennsylvania and University of Arizona.

Participants were administered both the PVT-192 and PVT-Touch during a visit to the laboratory during the daytime. Each test was 10 minutes long. Tests were presented in a block-randomized order, with N=5 receiving the PVT-192 first and N=5 receiving the PVT-Touch first, for each block of 10 participants. Thus, the presentation order was randomized but balanced across the sample. After each test presentation, subjects were asked to rate their distraction level, and the test administrator recorded any distracting noises or other issues.

We conducted Pearson correlations for mean RT, median RT, and number of lapses across subjects. To determine precision of the estimate, the Pearson *r* values were converted to Fisher's *z'* values, which are normally distributed. Then 95% confidence intervals were computed around the transformed *z'* values, then converted back to Pearson *r* values. To determine whether mean values from tests were systematically higher or lower from each other, paired *t*-tests evaluated differences between the 2 variables.

Due to device errors (devices losing power during the test, devices failing to initialize properly, and corrupted data), N=44 subjects completed both the PVT-Touch and the PVT-192 in a randomized order (N=6 hardware or software problems with the PVT-Touch, N=10 PVT-192 device failure, and N=3 noted distractions; some with multiple problems). PVT-Touch and PVT-192 correlated highly despite differences in input style and device hardware (Table 1). For mean RT, median RT, and lapses, correlations were all $r>0.8$. The PVT-Touch systematically produced slower reaction time values, and a greater number responses that would be scored as lapses (Table 1). To adjust for this difference, a scaling factor of 0.75 was applied to PVT-Touch mean and median RT. This resulted in no differences between groups, with t-test p values of 0.9994 for mean and 0.9526 for median. For lapses, a scaling factor of 0.5 was applied, resulting in a t-test p-value of 0.8617.

In-laboratory PVT scores have been shown to depend on time of day, duration of wakefulness, and can be influenced by a large number of environmental factors. Importantly, apps such as the PVT-Touch are typically not intended to replace the in-laboratory versions of the PVT, but rather to expand the research applicability of the PVT to more generalizable contexts than can be reasonably assessed with standard PVT approaches. With this in mind, a minimal validation of these apps may not necessitate exhaustively comparing results to the standard PVT across a wide range of circadian, sleep homeostatic, and environmental conditions – rather, a simple comparison of the two devices, demonstrating some degree of relative equivalence, should be sufficient as a first step. As more data becomes available, it can further clarify the degree to which a novel implementation of a PVT could generate useful data. Many commercially-available PVT products, though, fail to even meet this minimal standard. Either they are validated on different platforms (e.g., validation on desktop despite the commercial product being on a touchscreen device) or frequently not validated at all. It may be difficult to discern which PVT products are validated on the platform on which they are offered (and to what degree), whether they are validated on an alternate platform, or whether they are validated at all. This information should be transparently available.

In conclusion, consumer sleep technologies are changing the manner in which the public and the sleep research community conceptualize sleep assessment in the “real world”¹⁶. From the consumer perspective, these technologies allow longitudinal objective assessment of sleep in a typical sleep environment and individual assessment of the impact of behavior change on aspects of sleep. PVT-Touch contributes to this paradigm shift in sleep assessment by providing the opportunity to demonstrate changes in psychomotor vigilance linked to habitual sleep behaviors and can help guide individuals on their journey to optimal sleep health. The development and rigorous validation of touchscreen PVTs will enhance the sleep research community. First, it will reduce costs by increasing portability and accessibility. Second, sleep health researchers could use this type of software in large samples at low cost. Third, the general public, especially those with consumer wearables, may employ the software to assess their own performance. The sleep research community should work together to develop and welcome a new validated touchscreen based PVT. Preliminary results from PVT-Touch that our team has developed shows promise in achieving that goal.

ACKNOWLEDGEMENTS

This work was supported by a grant from the American Heart Association (12SDG9180007), the National Heart, Lung, and Blood Institute (K23HL110216), the National Institute for Environmental Health Sciences (R21ES022931), and the National Institute of Minority Health and Health Disparities (R01MD011600). Development of the PVT-Touch software was supported by the National Science Foundation (#IIS-1344613) and the Intel Science & Technology Center for Pervasive Computing. Dr. Kientz and Dr. Kay are supported by the National Science Foundation (#IIS-1344613) and the Intel Science & Technology Center for Pervasive Computing. Dr. Watson is supported by the National Cancer Institute (R03CA201806). Those interested in obtaining a copy of the PVT-Touch software should contact Dr. Grandner at grandner@email.arizona.edu or visit <http://pvttouch.com>.

REFERENCES

1. Dinges DF, Powell JW. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Beh Res Meth Instr Comp* 1985;17:652–5.
2. Lim J, Dinges DF. Sleep deprivation and vigilant attention. *Ann N Y Acad Sci* 2008;1129:305–22. [PubMed: 18591490]
3. Lo JC, Groeger JA, Santhi N, et al. Effects of partial and acute total sleep deprivation on performance across cognitive domains, individuals and circadian phase. *PLoS One* 2012;7:e45987. [PubMed: 23029352]
4. Basner M, Dinges DF. Maximizing sensitivity of the psychomotor vigilance test (PVT) to sleep loss. *Sleep* 2011;34:581–91. [PubMed: 21532951]
5. Zhou X, Ferguson SA, Matthews RW, et al. Mismatch between subjective alertness and objective performance under sleep restriction is greatest during the biological night. *J Sleep Res* 2012;21:40–9. [PubMed: 21564364]
6. Sunwoo BY, Jackson N, Maislin G, Gurubhagavatula I, George CF, Pack AI. Reliability of a single objective measure in assessing sleepiness. *Sleep* 2012;35:149–58. [PubMed: 22215929]
7. Basner M, Dinges DF. An Adaptive-Duration Version of the PVT Accurately Tracks Changes in Psychomotor Vigilance Induced by Sleep Restriction. *Sleep* 2012;35:193–202. [PubMed: 22294809]
8. Basner M, Mollicone D, Dinges DF. Validity and Sensitivity of a Brief Psychomotor Vigilance Test (PVT-B) to Total and Partial Sleep Deprivation. *Acta Astronaut* 2011;69:949–59. [PubMed: 22025811]
9. Thorne DR, Johnson DE, Redmond DP, Sing HC, Belenky G, Shapiro JM. The Walter Reed palm-held psychomotor vigilance test. *Behav Res Methods* 2005;37:111–8. [PubMed: 16097350]
10. Arsintescu L, Kato KH, Cravalho PF, Feick NH, Stone LS, Flynn-Evans EE. Validation of a touchscreen psychomotor vigilance task. *Accid Anal Prev* 2017.
11. Brunet JF, Dagenais D, Therrien M, Gartenberg D, Forest G. Validation of sleep-2-Peak: A smartphone application that can detect fatigue-related changes in reaction times during sleep deprivation. *Behav Res Methods* 2017;49:1460–9. [PubMed: 27631990]
12. Honn KA, Riedy SM, Grant DA. Validation of a portable, touch-screen psychomotor vigilance test. *Aerospace medicine and human performance* 2015;86:428–34. [PubMed: 25945658]
13. Kay M, Grandner MA, Bauer J, Lang RA, Watson NF, Kientz J. Initial validation of an Android-based psychomotor vigilance task. *SLEEP* 2013;36:A108.
14. Kay M, Rector K, Consolvo S, et al. PVT-Touch: Adapting reaction time test for touchscreen devices. In: *Proceedings of Pervasive Health 2013: The 7th Conference on Pervasive Computing Technologies for Healthcare*; 2013; Philadelphia, PA, 2013.
15. Smith A, Page D. US Smartphone Use in 2015: Pew Research Center, 2015 4, 2015.
16. Ko PR, Kientz JA, Choe EK, Kay M, Landis CA, Watson NF. Consumer Sleep Technologies: A Review of the Landscape. *J Clin Sleep Med* 2015;11:1455–61. [PubMed: 26156958]

Mean and Median Reaction Time and Attentional Lapses using the PVT-Touch and the PVT-192

Table 1.

	PVT-192		PVT-Touch		PVT-Touch (Corrected)		Correlation		
	Mean	SD	Mean	SD	Mean	SD	r	95% CI	p
Mean Reaction Time (ms)	284.10	54.83	378.81	98.13	284.09	73.57	0.83	(0.71, 0.91)	<0.0001
Median Reaction Time (ms)	258.75	39.00	344.50	68.44	258.17	51.34	0.82	(0.70, 0.90)	<0.0001
Lapses (>500ms)	4.32	5.85	8.20	11.48	4.10	5.74	0.83	(0.71, 0.91)	<0.0001