Acta Astronautica 69 (2011) 949–959

Contents lists available at ScienceDirect

Acta Astronautica

journal homepage: www.elsevier.com/locate/actaastro

Validity and sensitivity of a brief psychomotor vigilance test (PVT-B) to total and partial sleep deprivation

Mathias Basner [a,*], Daniel Mollicone [a,b], David F. Dinges [a]

[a] Unit of Experimental Psychiatry, Division of Sleep and Chronobiology, Department of Psychiatry, University of Pennsylvania, Perelman School of Medicine, 1013 Blockley Hall 423 Guardian Drive Philadelphia, PA 19104–6021, USA
[b] Pulsar Informatics, Inc., Philadelphia, PA, USA

ARTICLE INFO

Article history:
Received 7 December 2010
Received in revised form
27 May 2011
Accepted 7 July 2011
Available online 7 August 2011

Keywords:
PVT
Psychomotor vigilance
Alertness
Sleep deprivation
Sleep restriction
Attention
Lapse
Response time

ABSTRACT

The Psychomotor Vigilance Test (PVT) objectively assesses fatigue-related changes in alertness associated with sleep loss, extended wakefulness, circadian misalignment, and time on task. The standard 10-min PVT is often considered impractical in applied contexts. To address this limitation, we developed a modified brief 3-min version of the PVT (PVT-B). The PVT-B was validated in controlled laboratory studies with 74 healthy subjects (34 female, aged 22–45 years) that participated either in a total sleep deprivation (TSD) study involving 33 h awake ($N=31$ subjects) or in a partial sleep deprivation (PSD) protocol involving 5 consecutive nights of 4 h time in bed ($N=43$ subjects). PVT and PVT-B were performed regularly during wakefulness. Effect sizes of 5 key PVT outcomes were larger for TSD than PSD and larger for PVT than for PVT-B for all outcomes. Effect size was largest for response speed (reciprocal response time) for both the PVT-B and the PVT in both TSD and PSD. According to Cohen's criteria, effect sizes for the PVT-B were still large (TSD) or medium to large (PSD, except for fastest 10% RT). Compared to the 70% decrease in test duration the 22.7% (range 6.9–67.8%) average decrease in effect size was deemed an acceptable trade-off between duration and sensitivity. Overall, PVT-B performance had faster response times, more false starts and fewer lapses than PVT performance (all $p<0.01$). After reducing the lapse threshold from 500 to 355 ms for PVT-B, mixed model ANOVAs indicated no differential sensitivity to sleep loss between PVT-B and PVT for all outcome variables (all $P>0.15$) but the fastest 10% response times during PSD ($P<0.001$), and effect sizes increased from 1.38 to 1.49 (TSD) and 0.65 to 0.76 (PSD), respectively. In conclusion, PVT-B tracked standard 10-min PVT performance throughout both TSD and PSD, and yielded medium to large effect sizes. PVT-B may be a useful tool for assessing behavioral alertness in settings where the duration of the 10-min PVT is considered impractical, although further validation in applied settings is needed.

1. Introduction

Undisturbed sleep of sufficient length on a regular basis is of paramount importance for recuperation and the maintenance of behavioral alertness and cognitive performance [1,2]. Nevertheless, large parts of the population engage in acute or chronic partial sleep loss, suggesting that sleep is perceived as a flexible commodity that can be exchanged for waking activities considered more essential or of greater value [3]. In a recent analysis of time use in the US [4], work time was the waking activity most strongly reciprocally related to sleep time. At the same time the prevalence of shift work, requiring employees to both work and sleep at adverse times relative to their circadian phase, has increased

* Corresponding author. Tel.: +1 215 573 5866; fax: +1 215 573 6410.
E-mail address: basner@mail.med.upenn.edu (M. Basner).

0094-5765/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.actaastro.2011.07.015

over the past years [5]. Therefore, sleep disorders, lifestyle and work related curtailments of sleep, and working during unfavorable circadian times all may reduce neurobehavioral alertness to levels that increase the risk of errors and accidents [6,7]. Prevention of these outcomes through detection of fatigue (i.e., loss of alertness, sleepiness) remains a high priority in many safety-sensitive areas of human activity, and is also crucial for mission success in space flight.

Objective and quantitative assessments are necessary to evaluate the presence of fatigue-related deficits and to develop strategies for fatigue mitigation, especially as self-reports of sleepiness and self-assessments of performance capability have been shown to be unreliable [8,9]. In this context, neurobehavioral tests for fatigue assessment not only need to be operationally and conceptually valid, reliable, sensitive, specific, generalizable, and easy to use [10,11], but also brief enough to be acceptable for the target population and to allow for repeated administration in operational environments.

Many performance tests have been developed to objectively assess the degree of cognitive performance deterioration related to sleep loss. Among these, the Psychomotor Vigilance Test (PVT) is widely used [12,13]. It is based on simple reaction time (RT) to stimuli that occur at random intervals and therefore measures vigilant attention [14]. Auditory and visual reaction time tests have been used since the late 19th century in sleep research [15], but the PVT in its current version (i.e., 10-min duration with random inter-stimulus intervals (ISI) between 2 and 10 s) was proposed by Dinges and Powell in 1985 [16]. When appropriate PVT outcomes are used with precision timing of RT, the standard 10-min PVT has proven to be very sensitive to the dynamics of acute total sleep deprivation (TSD) and chronic partial sleep deprivation (PSD) [12].

Sleep deprivation causes both an overall slowing of PVT response times and an increase in the number of PVT errors of omission (i.e. lapses, usually defined as RTs $\geq$ 500 ms), as well as a smaller increase in errors of commission (responses without a stimulus) [14,17]. These effects increase with time on task [18]. An advantage the PVT has over nearly all other cognitive tests is that it is virtually unaffected by either aptitude (inter-individual variability) or learning (intra-subject variability)—that is, PVT performance does not improve as a function of repeated administration [19]. The test has high reliability, with intra-class correlations measuring test-retest reliability above 0.8 [13].

The 10-min PVT has been shown to be a valid tool for assessing behavioral alertness and vigilant attention performance in a large number of experimental, clinical, and operational paradigms. In addition to being sensitive to both TSD [17,20] and PSD [21,22], the PVT has demonstrated sensitivity to other perturbations of sleep homeostatic and circadian drives; [23,24] to inter- and intra-subject variability in the response to sleep loss; [9] to the effects of jet lag and shift work; [25] and to improvements in alertness following initiation of CPAP treatment in obstructive sleep apnea (OSA) patients; [26] administration of wake-promoting drugs; [27,28] and following naps. [29] Balkin et al. [30] assessed the utility of a variety of instruments for monitoring sleepiness-related performance decrements and concluded that the PVT "was among the most sensitive to sleep restriction, was among the most reliable with no evidence of learning over repeated administrations, and possesses characteristics that make it among the most practical for use in the operational environment."

The standard 10-min PVT with 2–10 s ISI is most commonly used, although both longer [18,31] and shorter [32] duration versions have been evaluated. Test duration is an important aspect of the PVT because even severely sleep deprived subjects may be able to perform normally for a short time by increasing compensatory effort. However, in a systematic analysis of PVT duration, we showed that the ability of the PVT to differentiate alert and sleepy subjects was, depending on the outcome variable, only marginally lower (and at times higher) for shorter than 10-min test durations [12]. Therefore, optimal PVT duration may be shorter than 10-min for some outcome variables, demonstrating feasibility of shorter versions of the PVT. Accordingly, a 5-min handheld version of the PVT already exists [32–36]. However, both 2-min [32] and 90 s [34] versions of the PVT were deemed to be too insensitive to be used as valid tools for the detection of neurobehavioral effects of fatigue, leaving open the question of whether a brief PVT that was sensitive to sleep loss could be developed.

We therefore set out to develop a brief PVT (PVT-B) that was as sensitive to TSD and PSD as the standard 10-min PVT. Based on our theory of how sleepiness manifests in performance, our large PVT databases, knowledge on the importance of outcome variable [12], ISI, and precision of timing for the ability of the PVT to differentiate sleep deprived and alert subjects, we shortened test duration from 10 to 3 min and ISI from the standard 2–10 to 1–4 s to create the PVT-B, while maintaining sufficient response sampling rates to detect wake state instability. [14] We hypothesized that PVT-B would retain its sensitivity and specificity to sleep loss, and therefore be a practical tool for fatigue assessment. A sensitive, specific, brief PVT-B would meet the criteria for fitness-for-duty testing not only prior to the start of a shift but also during repeated administrations while on the shift.

2. Material and methods

2.1. Subjects and protocol

This investigation used data from a TSD and from a PSD protocol. The TSD data were gathered in a study on the effects of night work and sleep loss on threat detection performance on a simulated luggage screening task (SLST). A detailed description of the study is published elsewhere [37]. This analysis is based on data gathered in a pilot study on $N=12$ subjects and in the main study on $N=24$ subjects. Four subjects were excluded from the analysis due to non-compliance or excessive fatigue during the first 16 h of wakefulness. Another subject withdrew after 26 h awake. Therefore, a subset of $N=31$ subjects (mean age $\pm$ standard deviation $=31.1 \pm 7.3$ yr, 18 female) contributed to the analyses. Study participants stayed in the research lab for five consecutive days, which included a 33 h period of TSD.