



Functional Equivalence of Sleep Loss and Time on Task Effects in Sustained Attention

Bella Z. Veksler, Glenn Gunzelmann

Air Force Research Laboratory, Wright-Patterson Air Force Base

Received 27 April 2015; received in revised form 2 December 2016; accepted 15 December 2016

Abstract

Research on sleep loss and vigilance both focus on declines in cognitive performance, but theoretical accounts have developed largely in parallel in these two areas. In addition, computational instantiations of theoretical accounts are rare. The current work uses computational modeling to explore whether the same mechanisms can account for the effects of both sleep loss and time on task on performance. A classic task used in the sleep deprivation literature, the Psychomotor Vigilance Test (PVT), was extended from the typical 10-min duration to 35 min, to make the task similar in duration to traditional vigilance tasks. A computational cognitive model demonstrated that the effects of time on task in the PVT were equivalent to those observed with sleep loss. Subsequently, the same mechanisms were applied to a more traditional vigilance task—the Mackworth Clock Task—providing a good fit to existing data. This supports the hypothesis that these different types of fatigue may produce functionally equivalent declines in performance.

Keywords: Vigilance; Fatigue; ACT-R; Cognitive modeling; Time on task; Psychomotor vigilance

1. Introduction

Sustaining attention for prolonged periods of time is crucial in many tasks in both the workplace and in school, whether you are driving a car, working as a baggage screener, or simply attempting to pay attention to a lecture. These kinds of tasks—often called sustained attention or vigilance tasks—are well-known for leading to declining performance (referred to as the vigilance decrement) as time on task increases (e.g., Davies & Parasuraman, 1982). While there have been many theories put forth to account for the vigilance decrement, computational accounts are lacking. The goal of this work is to test the hypothesis that cognitive declines associated with the vigilance decrement are

functionally equivalent to fluctuations in cognitive performance related to time awake and circadian rhythms. We do this by evaluating the capacity of a computational theory of sleep loss to account for the vigilance decrement.

Mackworth (1948) described vigilance as the “psychological readiness to perceive and respond” (p. 133). Being vigilant, therefore, manifests itself behaviorally as fast and accurate responses to stimuli. Traditionally, vigilance tasks require participants to monitor a stream of stimuli and respond when a rare signal occurs. Human performance on this task declines systematically with increasing time on task. In particular, there is a decrease in detection accuracy and a corresponding increase in response time, which is most pronounced during the first 30–45 min of the task (Mackworth, 1948). The vigilance decrement has been demonstrated in a variety of contexts, including visual detection (Adams & Boulter, 1964; Bakan, 1955; Hitchcock et al., 2003; Jerison & Pickett, 1964; Wiener, Curry, & Faustina, 1984; Williges, 1973), auditory detection (Kennedy & Coulter, 1975; Levine, 1966; Loeb & Binford, 1964), and tasks varying greatly in complexity (see Molloy & Parasuraman, 1996, for a review).

The foundation of our computational account of the causes of the vigilance decrement is an existing theory and set of mechanisms developed to explain fluctuations in performance related to sleep loss and circadian rhythms (Gunzelmann, Gross, Gluck, & Dinges, 2009; Walsh, Gunzelmann, & Van Dongen, 2014). Although theories about the impact of sleep loss and theories of the vigilance decrement are largely distinct (but see Gunzelmann, Moore, Gluck, Van Dongen, & Dinges, 2010; Krueger et al., 2008), it is our goal to highlight connections between these areas of study, culminating in a proposal that the same mechanisms can account for the changes in performance that result from both sleep loss and prolonged time on task. We demonstrate that, although the degradations operate at different time scales (hours and days in the case of sleep deprivation and minutes to hours in the case of vigilance studies), they are functionally equivalent in terms of their impact on cognitive processing. In both cases, our model suggests that the resources necessary for efficient and effective information processing by the cognitive system are depleted. Analogous recent findings from neuroscience suggest that sleep occurs at both the neuronal assembly and behavioral levels (Krueger et al., 2008; Van Dongen, Belenky, & Krueger, 2011). Furthermore, instances of “sleep” occurring at the more local neural level may depend on prior activity within the network (Krueger et al., 2008).

2. Study of vigilance

Over the past century, there has been considerable evidence in the literature that perceptual-motor detection and discrimination tasks suffer from a vigilance decrement (Bakan, 1955; Mackworth, 1948; Parasuraman, 1979). While other types of performance decrements have been studied as well, for example, the speed of writing and reading sequences of letters (Robinson & Bills, 1926), the focus in this manuscript will be specifically on tasks requiring perceptual detection of a signal. Typical vigilance tasks contain rare signals that require a response interspersed with more frequent stimuli that must be

ignored. A classic study by Mackworth (1948) illustrates the type of task generally used to study the vigilance decrement. In the Mackworth Clock Task, participants monitor the forward ticks of a clock hand and respond when the tick is twice the usual magnitude.

Vigilance decrements have also been studied with tasks such as determining when a stimulus on a radar display changes (Adams & Boulter, 1960), detecting luminance changes (Bakan, 1955), or identifying an imminent collision on simulated radar displays (Hitchcock et al., 1999). The common thread through all of these tasks is that individuals exhibit progressively lower probabilities of target detection and slower reaction times with longer time on task. While a vigilance decrement has been found in a varied range of tasks throughout its long history of research, the magnitude of the decrement has been shown to be modulated by various factors, including event rate, signal probability, stimulus duration, presence of other tasks, stimulus modality, and expectancy, among others (Bevan, Avant, & Lankford, 1967; Smith, Lucaccini, & Epstein, 1967; Smith, Lucaccini, Groth, & Lyman, 1966). These factors highlight important nuances of the decrement and provide a set of criteria for developing and extending theory. For purposes of the current paper, however, we focus on a basic mechanism that produces a decrement, providing an initial computational theory of the core phenomena.

Currently, the most widely accepted theory ("Resource Theory") associates the vigilance decrement with the depletion of information-processing resources that cannot be replenished adequately during task performance (Parasuraman, 1979; Warm, Parasuraman, & Matthews, 2008). Support for this view comes from studies in which workload during a vigilance task is measured using the NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988). Workload scores increase with time on task mirroring performance declines, especially in tasks where memory is taxed, which has led researchers to conclude that *Vigilance Requires Hard Mental Work and Is Stressful* (Warm et al., 2008). The resource account is useful in describing some of the behavioral changes observed in vigilance tasks.

In this paper, we describe a computational theory of the vigilance decrement that instantiates a quantitative version of Resource Theory and provides a detailed explanation of the underlying changes in cognitive processing that lead to declines in performance during sustained attention tasks. Meanwhile, the changes in model performance that are brought on by resource depletion are consistent with an alternative theory (*mind wandering*, e.g., McVay & Kane, 2010) and a more recent integrative proposal (e.g., Thomson, Besner, & Smilek, 2015), potentially helping to unify opposing theoretical accounts of the vigilance decrement. The details of the model and its relation to theories of the vigilance decrement are discussed in detail below.

In addition to extensive empirical evidence, there has been descriptive mathematical modeling of the vigilance decrement, characterizing how performance declines with increasing time on task (Fisk & Schneider, 1981; Giambra & Quilter, 1987; Mackworth, 1964; Teichner, 1974). These models have found that the vigilance decrement is best fit by an exponential curve, which is characterized by a rapid initial drop and subsequent asymptote. It is the goal of the current work to associate those dynamics with underlying changes to information processing in the cognitive system that lead to the performance

decrement, and to evaluate whether those mechanisms match the ones that have been previously validated in the context of sleep loss (e.g., Gunzelmann, Gross et al., 2009). This will broaden the theory and provide a more detailed theoretical account of the vigilance decrement, which has the potential to generalize to other tasks and dependent measures.

In the remainder of the paper, we will discuss the theoretical background for the computational mechanisms, our approach for implementation, and a description of a novel experiment that helps to bridge the literature on sleep loss and time on task. We will follow this up with descriptions of the models and their comparison to the human data.

3. Fatigue theory

In the same way that mathematical modeling has been used to model the vigilance decrement, there have also been attempts to develop mathematical models to capture the deleterious effects of sleep loss and circadian rhythms on cognitive performance (e.g., Mallis, Mejdal, Nguyen, & Dinges, 2004). These models, referred to as biomathematical models, are used to estimate cognitive performance levels based upon systems of equations representing time awake and circadian rhythms, and sometimes additional factors that influence dynamic fluctuations in fatigue (e.g., light exposure—Duffy, Kronauer, & Czeisler, 1996; sleep inertia—Jewett & Kronauer, 1999; caffeine—Ramakrishnan et al., 2014). They have proven effective in capturing the overall dynamics of these influences with sleep deprivation (Van Dongen, 2004), and more recently with sleep restriction (McCauley et al., 2013).

A limiting factor in biomathematical models is that they do not directly make predictions about either performance or the underlying changes in cognitive processing (Gunzelmann, Veksler, Walsh, & Gluck, 2015). Instead, they produce estimates of overall cognitive functioning, which we refer to as *alertness*. The outputs can be fit to specific dependent measures, but this requires a data-driven scaling process (Van Dongen, 2004). At a theoretical level, they do not address the mechanisms that produce behavior and performance changes.

Validation of biomathematical models of alertness often compares model outputs to performance on a reaction time task called the Psychomotor Vigilance Test, or PVT (Dinges & Powell, 1985). The PVT is used extensively in sleep research, in large part due to the sensitivity of performance to time awake and circadian rhythms (e.g., Lim & Dinges, 2008). In the task, participants monitor a known location on a monitor and respond when a stimulus appears by pressing a response button. The stimulus is an incrementing counter that shows the number of milliseconds that have elapsed since the stimulus appeared. The task taxes sustained attention by manipulating the inter-stimulus interval. Stimuli are presented at intervals ranging from 2 to 10 s following the response to the previous stimulus. With sleep deprivation, response time distributions become increasingly positively skewed, and more responses are characterized as “lapses” (response times greater than 500 ms). Biomathematical model predictions are usually scaled to the number of lapses observed in participants during a 10-min PVT

session (Van Dongen, 2004). With increasing sleep loss, the occurrence of lapses increases dramatically, modulated by circadian rhythms (see Lim & Dinges, 2010, for a review).

Gunzelmann, Gross et al. (2009) proposed a set of mechanisms to represent the impact of fluctuations of alertness on information-processing activity (Walsh et al., 2014). While theories of sleep loss have referred to concepts like cognitive slowing and cognitive lapsing (Dinges & Kribbs, 1991; Kleitman, 1963), Gunzelmann, Gross et al. (2009) were able to demonstrate the capacity of a computational model to produce changes in performance that closely matched changes in human behavior across 88 h of total sleep deprivation. The critical mechanism in the model introduced brief disruptions in goal-directed cognitive processing, termed *microlapses*. Because the disruptions were brief (10's of milliseconds), performance was frequently slowed, reproducing the characteristic slowing of response times reflected in increasingly right-skewed response time distributions in the PVT. However, *microlapses* also increased in probability with increased time awake, leading to more significant performance lapses.

Microlapses implement a version of the state-instability hypothesis, which has been proposed as an explanation for performance breakdowns associated with sleep loss (Doran, Van Dongen, & Dinges, 2001). According to this theory, sleep loss leads to a fundamentally unstable cognitive state in which the cognitive system vacillates between sleep and wakefulness. Without effort, the system rapidly transitions to sleep. Even with effort, however, performance degrades with decreasing alertness as the pressure to sleep overwhelms the attempts to maintain wakefulness. Because of the duration of a microlapse, and the potential for sequences to produce more substantial performance disruptions, microlapses accounted for both, the slowing of responses and the increase in lapses that are characteristic changes in the PVT associated with sleep loss and circadian rhythms (Gunzelmann, Gross et al., 2009).

In addition to accounting for the negative impact of fatigue, Gunzelmann, Gross et al. (2009), also incorporated a mechanism reflecting the impact of compensatory effort. In the model, compensation was implemented as a manipulation to a threshold parameter that influenced the likelihood of a *microlapse* (more details are provided below). The mechanism decreased the probability of a *microlapse*, in effect, by diminishing inhibition in the execution of cognitive actions (c.f., Drummond, Paulus, & Tapert, 2006). While this led to a decrease in *microlapses*, it also resulted in diminished inhibition more generally, which produced a corresponding increase in errors of commission due to executing a response when no stimuli had yet appeared (i.e., false starts). This behavior aligns with observations of human performance when sleep deprived, and with neuropsychological data related to the relationship between effort and activation in the thalamus in sustaining performance during sleep deprivation (e.g., Chuah, Venkatraman, Dinges, & Chee, 2006; Portas et al., 1998).

Overall, the mechanisms provided a detailed account of how sleep loss and circadian rhythms impact PVT performance, with the behavior of the model closely aligned to human performance. A critical component of the research included theoretically linking parameters in the model to the dynamics of alertness predicted by the biomathematical

models described earlier. Subsequent research has illustrated how the same mechanisms can be used to generate performance predictions in other contexts, including dual tasking (Gunzelmann, Byrne et al., 2009) and driving (Gunzelmann, Moore, Salvucci, & Gluck, 2011). The implementation of this theory is described next.

3.1. Framework for theory implementation

The theory was tested by implementing it in a cognitive architecture. Cognitive architectures, based on decades of empirical evidence and psychological theory, have successfully reproduced many effects found in the psychology literature (e.g., Anderson, 2007; Laird, 2012). One benefit of using a cognitive architecture is that it provides a unified theory of a myriad of cognitive constructs all working together and providing the most complete theory of human cognition (Newell, 1990). This facilitates cumulation and generalizability. In the case of this research, the cognitive architecture provides an existing computational theory that can be leveraged for evaluating mechanisms to explain phenomena in a new domain of interest, namely fatigue.

To model human performance on the PVT, we used Adaptive Control of Thought-Rational, or ACT-R (Anderson, 2007). ACT-R is a cognitive architecture which instantiates a theory of the mechanisms of human cognition. ACT-R is a modularized production system of if/then rules, which serve as its central cognition. These production rules interface with other components of cognition, such as visual and motor modules, through limited capacity buffers that allow information to be passed in both directions. ACT-R's production system uses a utility mechanism to determine which productions should be executed and generates behavioral predictions based on cognitive constraints that have been validated over decades of research. Additionally, it is capable of interfacing with the same environments as human participants making it an excellent framework for evaluating theoretical accounts of many phenomena, including the vigilance decrement. Hundreds of models have been developed in ACT-R, which provide quantitative accounts for a wide variety of phenomena (Anderson, 2007).¹

Central cognition within ACT-R operates through cycles where candidate productions are identified based on their match to the current context (buffer contents); one is selected and (usually) executed. The process is governed by utility values associated with each production. During the production selection process, a utility value is computed for each production based upon its history of use, reinforcement learning, and match to the current context (i.e., contents of the buffers, including representations of what is in the environment). All productions whose preconditions match the state of the cognitive system (i.e., contents of buffers) are placed in a *conflict set*. To be selected, a production must have the highest utility value (U_i , out of the *conflict set*), and that value must exceed the utility threshold (UT). Typically, the utility threshold remains static throughout model execution while production utilities are recalculated during each production cycle and compared both to each other and to the UT . Noise in utility values produces stochasticity in model behavior.

3.2. Fatigue mechanisms

One of the benefits of using ACT-R as the modeling framework is that it is straightforward to extend its functionality by adding additional modules to the architecture to extend its explanatory power. Our previous work within the ACT-R architecture has resulted in the addition of a fatigue module to make predictions about performance degradations as a result of sleep deprivation (Gunzelmann, Gross et al., 2009). Unlike most modules in ACT-R, the fatigue module does not contain mechanisms for processing information, nor a buffer to make the results of its processing available to central cognition. Instead, the module acts as a moderator of other modules in the architecture, influencing the efficiency and effectiveness of cognitive processing in other parts of the system by influencing parameter values associated with their mechanisms. For the current model, the module modifies three key components of the architecture: the behavior of the model when no productions have utilities above threshold, utility calculations, and the value of the utility threshold. This is consistent with evidence regarding the role of neural structures associated with sleep and circadian processes in the human brain (e.g., Saper, Chou, & Scammell, 2001). Moreover, the implementation of the module is also consistent with findings that subjective reports of fatigue tend to be relatively inaccurate (Leproult et al., 2003), suggesting that individuals do not have access to information about the state of the alertness system.

ACT-R's default behavior is to halt model execution if no productions have utilities that rise above the utility threshold. In most ACT-R models, this is used to signal the attainment of a goal (i.e., end of a trial). Occasionally, this has been used to represent "giving up" in models of motivation (Belavkin, Ritter, & Elliman, 1999). In contrast to this default behavior, the fatigue module allows model execution to proceed even after a production cycle in which no production utilities exceed the utility threshold. When this happens the production cycle, lasting 50 ms on average in ACT-R by default, passes with no cognitive actions being performed (i.e., a *microlapse*, formally defined in Eq. 3 below), creating a small gap in cognitive processing equivalent in duration to a single production cycle (10's of ms). This represents the primary consequence of resource depletion resulting from fatigue, where *utility* represents the resource. To allow for the model to terminate execution, the modeler now specifies a model run duration equivalent to the task duration.

Following Gunzelmann, Gross et al. (2009), the fatigue mechanism modifies the utility value of all productions which match the current context as follows:²

$$\text{Utility}(t) = FP(t) \times U_i(t) + \epsilon \quad (1)$$

where $U_i(t)$ is the non-fatigued utility value of production i at time t and $FP(t)$ is the fatigue modulation value at time t (constrained between 0 and 1), derived from biomathematical model predictions of alertness (defined formally below). Logistically distributed noise is subsequently added to produce stochasticity in model performance, denoted as ϵ . Importantly, fatigue is theorized to impact utility values themselves, but not the noise that

influences them on a given cognitive cycle. Because fatigue scales utility values proportionally, this increases the influence of noise on production selection with greater fatigue.

In Gunzelmann, Gross et al. (2009), the fatigue module was validated with participants whose sleep histories were carefully controlled and who performed the same tasks in numerous sessions over the course of several days of sleep deprivation. However, participants in this study only performed the PVT once, and self-reported their sleep histories. Because of the lack of data to drive biomathematical model predictions of alertness, we omitted that component of the FP calculation. Variability in initial performance stemming from sleep history, motivation, and other factors is absorbed into estimates of the Initial Utility ($U_i(0)$) for each individual³ in the results presented below. For purposes of this paper, $FP(t)$ is based upon the time-on-task component and a decrement due to *microlapses* during the trial (FP_{pct} from Eq. 3 below):

$$FP(t) = FP_{pct} \times (1 + t)^{PU_{TOT}} \quad (2)$$

where t is the amount of time (in minutes) since the task began and PU_{TOT} is a parameter that reflects the production utility time-on-task decline (the range of possible values for PU_{TOT} is $[-1,0]$). At the beginning of task execution, $FP(t)$ is 1 and declines as time on task elapses, which causes production utilities to decrease. This equation reflects mathematical characterizations of the vigilance decrement, which has fit exponential functions to vigilance data (Giambra & Quilter, 1987).

The probability of *microlapses* is influenced by the utility values of the productions and the utility threshold. The fatigue module produces reductions in both values as alertness declines. Lower utility values lead to more *microlapses*. Additionally, within each trial, each subsequent *microlapse* further degrades the fatigued utility calculation. This exacerbates the decline in production utilities and contributes to the skewedness of the response time distribution produced by the model of the PVT. Formally, the FP_{pct} is recalculated after each *microlapse* as follows:

$$FP_{pct} = (FP_{dec})^{microlapses} \quad (3)$$

For the current work we used the same degradation value ($FP_{dec} = 0.98$) following a *microlapse* as in Walsh et al. (2014). Stochasticity in utility values makes it possible, often likely, that a *microlapse* will be followed by a production cycle where appropriate actions are executed. In these cases, disruptions produce modest slowdowns in performance. However, *microlapses* can accumulate to produce substantial changes in cognitive performance and behavior. This is particularly true in time-critical tasks, where even small slowdowns can have significant consequences for performance (e.g., Gartenberg, Veksler, Gunzelmann, & Trafton, 2014).

In addition to the changes in production utility, the fatigue module also alters the utility threshold. The utility threshold dictates the minimum utility value a production must have to be selected. Without the fatigue module, this is set as a constant parameter and does not change with model execution. To offset the negative consequences of degraded

utility values, a compensatory mechanism lowers the utility threshold in concert with degraded alertness, which reduces the probability of *microlapses*. This can be interpreted as “effort” and reflects evidence for increased workload over time in vigilance tasks (e.g., Warm et al., 2008) and in neuropsychological research linking increased activation in the thalamus with increased effort and performance maintenance in sleep-deprived individuals (Portas et al., 1998). Reducing the threshold, however, also has the side effect of increasing the probability of executing productions with lower utilities, including productions that only partially match the current context. The consequence is that stochasticity plays an increasing role in model execution as alertness declines, leading to a higher likelihood of both errors of commission and errors of omission (c.f., Doran et al., 2001). The decrease in utility threshold is not affected by within-trial dynamics (*microlapses*). Instead, the fatigue module modifies the utility threshold (to represent increased effort as compensation for feeling tired) in a manner similar to the production utility modulations:

$$\text{Utility Threshold}(t) = FT(t) \times UT(0) \quad (4)$$

where $UT(0)$ is the Initial Utility Threshold set by the modeler. The setting of $UT(0)$ is intended to capture motivational factors (e.g., interest in the task, or desire to perform well in the experiment). $FT(t)$ is the fatigue module’s modulation value calculated as:

$$FT(t) = (1 + t)^{UT_{TOT}} \quad (5)$$

where UT_{TOT} is the utility threshold time-on-task decline parameter. This mechanism, coupled with the theoretical mapping of ACT-R to brain regions, is consistent with the evidence regarding the role of effort in offsetting the negative effects of sleep loss (e.g., Chuah et al., 2006; Doran et al., 2001; Drummond et al., 2000).

Taken together, production utilities for all productions which match based on their buffer contents are calculated using Eq. 1 and are then compared with the Utility Threshold from Eq. 4 in order to determine which, if any, production should fire during a particular cycle. In the context of both models in the current work, determination of whether a production’s buffer contents match the state of the world relies on the production partial matching mechanism described below.

3.3. Production partial matching

The fatigue mechanism just described provides a way for model performance to degrade; however, there is an additional ACT-R mechanism that influences production selection, which plays a critical role in the current research. *Production partial matching* allows productions to be selected even if the values in the buffers are not exactly those specified by the production. In the context of the theoretical account, this mechanism allows for the execution of inappropriate cognitive actions, which has been implicated in prior research (Walsh et al., 2014). In the model, it leads to errors of commission, observed as *false starts* in the PVT and *false alarms* in vigilance tasks.

A scaling parameter, *ppm*, dictates the degree of discounting of a production's utility value to the extent that the production's conditions do not match the current state (buffer contents). The equation that utilizes this parameter is:

$$U_i(t) = U_i(t)' + \sum_j^N ppm \times \text{similarity}(d_j^i, v_j^i) \quad (6)$$

where $U_i(t)'$ is production i 's current utility value prior to production partial matching, *ppm* is the scaling parameter, d_j^i is the desired value for slot j in production i , and v_j^i is the actual value for slot j in production i . The function computing the similarity between d_j^i and v_j^i is assumed to be 0 when the expected and actual values are identical, and a negative value when they differ. The utility value resulting from this calculation is then modulated by noise and by the fatigue mechanisms described in the previous section (Eq. 1).

The fatigue mechanisms, together with the production partial matching mechanism, modify model performance across time on task in the same manner as was done for sleep loss in Gunzelmann, Gross et al., 2009; Walsh et al., 2014. We next describe an experiment involving an extended PVT and model fits to the data. Subsequently, we report model fits to published data for a traditional vigilance task to demonstrate the generalizability of these mechanisms to time on task effects.

4. Current approach

Building upon the model in Gunzelmann, Gross et al. (2009), Gunzelmann et al. (2010) extended the model to account for changes in performance on the PVT that unfold over the course of a 10-min session. The model utilized the same mechanisms as the original model, but introduced a time-on-task component, which had impacts on model performance that were additive with respect to the effects of sleep loss and circadian rhythms. The current effort extends this prior research in systematic ways. First, performance on the 10-min PVT has been shown to be highly sensitive to fluctuations in alertness. However, the PVT has not been tested at durations that are more typical in the vigilance literature. We expect that an extended PVT task, lasting more than 30 min, will show a pattern of decline that is aligned with research on the vigilance decrement (i.e., an exponential decline with time on task). Further, we expect that the same mechanisms that accounted for performance degradations stemming from sleep loss and circadian rhythms will continue to produce performance changes in the model that are well aligned with changes in human performance associated with time on task.

In addition to extending the model of the PVT to longer sessions, we also test the generalizability of the mechanisms to a traditional vigilance task. For this, we use results from two published studies involving the Mackworth Clock Task (Giambra & Quilter, 1987; Mackworth, 1948). This is a critical test of the model due to systematic differences

between the PVT and tasks commonly used in research on the vigilance decrement. These differences are discussed after the PVT model results.

5. Experiment

The sleep research community has been using the Psychomotor Vigilance Test (PVT) for several decades to explore the effects of sleep restriction/deprivation on performance (Dinges & Powell, 1985; Lim & Dinges, 2008; Van Dongen & Dinges, 2005a). Participants in this task must respond as quickly as possible when an incrementing counter appears on the screen. Typically, the PVT lasts for 10 min (Dinges & Powell, 1985), though versions as short as 3 min (Basner & Dinges, 2011) and as long as 20 min (e.g., Van Dongen & Dinges, 2005b) have been investigated. In contrast, traditional vigilance tasks typically span more than 30 min (Davies & Parasuraman, 1982). Thus, we presented participants with a 35-min PVT to compare decrements associated with time on task to those that have been found for sleep loss.

5.1. Participants

A total of 20 University of Dayton undergraduates and local area participants completed the study (12 Female) with a mean age of 22.35 ($SD = 2.45$). Participants were paid \$15 for their participation. All participants had normal or corrected-to-normal vision. The protocol was approved by the Institutional Review Board of the Air Force Research Laboratory. All participants provided written informed consent.

5.2. Materials

A 17-inch monitor with the resolution set to $1,280 \times 1,024$ pixels displayed a gray background. Participants continuously monitored the computer screen for the appearance of an incrementing counter and responded by pressing a button on a button box as quickly as possible. The counter was a four-digit number indicating the time in milliseconds since its appearance. It was drawn in 80-point black Arial font in the middle of the screen and appeared after a uniformly random delay between 2 and 10 s from the previous trial's response. The response had to be made within 30 s of the counter appearing to be counted as a valid response, otherwise the trial timed out (termed a *sleep attack* or *non-response* in previous PVT literature). There were no *non-responses* observed in this study. Once a response was made, the counter stopped incrementing and displayed the participant's reaction time as feedback for 500 ms. The dependent measure of response time was recorded in milliseconds. Participants completed 340 trials. The entire task took about 35 minutes to complete (range: 33.17–37.27, $M = 35.77$, $SD = 0.95$). Participants were tested between the hours of 9 am and 5 pm, Monday through Friday.

5.3. Results

Prior to discussing the results of the study in detail, it should be noted that because PVT performance is sensitive to sleep loss, we asked participants about their sleep history prior to beginning the task (i.e., the number of hours slept the night before and the time they woke up that morning). Participants reported between 5 and 10 h of sleep the night before ($M = 7.63$, $SD = 1.31$) and woke up between 7:30 am and 12 pm ($M = 8:57$ am, $SD = 1.21$). To test whether time in bed (TIB) or wake up time (T0) had an effect on performance in this task, we first determined the difference in performance between the first and last block of 5 min of the task. We found no correlation between TIB and the change in median RT between the first and last block, $r = -.26$, $n = 20$, $p = .26$. We also found no correlation between TIB and the change in lapses between the first and last block, $r = -.38$, $n = 20$, $p = .1$. There was also no correlation between TIB and the change in false starts between the first and last block, $r = .19$, $n = 20$, $p = .42$.

We did the same analysis for wake up time and found no effect on median RT difference, $r = -.25$, $n = 20$, $p = .28$; no effect on lapses, $r = -.04$, $n = 20$, $p = .87$; and an effect on false starts, $r = .49$, $n = 20$, $p = .03$. This last effect was largely due to the higher false starts in two participants, both of whom woke up fairly late the day of their study (10:30 am and 11 am).

In addition to sleep, caffeine has been shown to alter performance on attention intensive tasks (Kilpeläinen, Huttunen, Lohi, & Lyytinen, 2010; Temple et al., 2000; Van Dongen et al., 2001), though these effects are generally not defined quantitatively in existing models of alertness (but see Benitez, Kamimori, Balkin, Greene, & Johnson, 2009; Ramakrishnan et al., 2013). Out of our sample of 20, five participants reported consuming coffee on the day of the study. We did not find a significant effect of coffee on the difference in median RT from block 1 to 7 $t(18) = -0.23$, $p = .82$, lapses $t(18) = -0.04$, $p = .97$, or false starts $t(18) = 1.49$, $p = .15$ and so it is not considered further in the analyses that follow.⁴

PVT performance is typically characterized using three or four dependent measures. These include the following:

1. *False Starts*: Button presses before the stimulus appears, or within 150 ms of stimulus onset (considered to be too fast to be an intentional response)
2. *Median Reaction Time*: Because the response time distribution for PVT performance has a positive skew, this measure provides a more stable index of central tendency than the mean.
3. *Lapses*: Responses > 500 ms. This measure is used extensively in the sleep research literature as an index of fatigue, since it has been shown to be highly sensitive to fluctuations in alertness associated with time awake and circadian rhythms.
4. *Sleep Attacks*: In some studies involving significant sleep deprivation (e.g., Dinges & Powell, 1985; Doran et al., 2001), participants will occasionally go more than 30 s without responding to a stimulus in the PVT. In these cases, the trial is terminated, and a new trial is initiated (participants are also reminded to focus on the

task and respond as quickly as possible). There were no sleep attacks observed in this study.

Importantly, the measures that are used to characterize PVT performance are derived from response time distributions produced by participants based upon responses to individual stimuli in the task. Historically, theories of fatigue have not been able to capture the full response time distribution, necessitating these alternative representations. Because our model does produce response time distributions that can be compared to human PVT data, we also present probability distributions and cumulative distributions to evaluate human performance in more detail. This also provides the foundation for a more detailed assessment of the model, which is presented later.

5.3.1. Median response time

For each participant, the median of all response times within each block, excluding false starts, was used in the analysis (the data were divided into seven 5-min blocks), in line with previous work showing median RT to be a reliable indicator of PVT performance (Basner & Dinges, 2011). We conducted a repeated-measures ANOVA with block as the repeated variable and found a significant difference in median response time as time-on-task increased, $F(6, 114) = 8.55, p < .001$. The median response times for the seven 5-min blocks are presented in Table 1.

5.3.2. Percent lapses

The percentage of responses characterized as lapses (response times longer than 500 ms) also increased across blocks, $F(6, 114) = 5.44, p < .001$. Table 1 contains the percentage of lapses for the seven 5-min blocks.

5.3.3. Percent false starts

There was not a significant change in mean percent false starts by 5-min blocks across time on task, $F(6, 114) = 1.46, p = .199$. While not significant, there appears to be a trend toward an increasing percentage of false starts over the seven 5-min blocks (see Table 1).

Table 1
Psychomotor Vigilance Test (PVT) results with 95% confidence intervals

5-Min Block	Response Time (ms)	% Lapses		% False Starts	
	Median	M (SD)	95% CI	M (SD)	95% CI
1	285	1.17 (1.71)	[0.42, 1.92]	0.64 (1.00)	[0.2, 1.07]
2	290	1.67 (2.05)	[0.77, 2.57]	0.31 (0.75)	[−0.02, 0.64]
3	293	1.75 (3.15)	[0.37, 3.13]	1.34 (2.98)	[−0.03, 2.64]
4	302	2.42 (2.43)	[1.35, 3.48]	1.58 (2.79)	[0.35, 2.8]
5	307	3.69 (4.58)	[1.68, 5.7]	1.74 (3.49)	[0.21, 3.27]
6	308	2.90 (2.72)	[1.71, 4.09]	1.55 (2.98)	[0.25, 2.86]
7	312	5.23 (4.81)	[3.12, 7.34]	2.13 (4.52)	[0.15, 4.12]

5.3.4. RT distribution

As a measure of central tendency, median response times provide a concise characterization of changes with time on task. However, a response time distribution gives us a more precise picture of how performance changes. Fig. 1 (left) presents probability density functions (PDFs) for each of the 5-min blocks in the data. As time on task increases, the response time distribution gradually stretches to the right, reflecting the slowing of responses and increase in lapses.

Although PDFs provide a helpful way of viewing the data, cumulative distribution functions (CDFs) offer some advantages for comparing human performance to model results. CDFs have been used in the past for fitting response time distributions, helping to account for the characteristic positive skew of response times (Van Zandt, 2000). The CDF specifies how much of the data in the distribution falls below each predetermined boundary. For purposes of these data, we used the data from all participants for the first 5-min block to determine the 5% quantile boundaries of the response time distribution for all responses made between 150 ms and 30 s (boundaries at 235, 243, 250, 255, 261, 266, 271, 275, 280, 285, 288, 294, 301, 307, 315, 325, 335, 352, 383, and 845 ms beginning at 5% in 5% increments). These boundaries were then used to determine how much of the distribution fell below each boundary in each subsequent block. We incorporated *false starts* into the CDF by including the percentage of *false starts* as an additional quantile (the leftmost data point at 0ms RT in Fig. 1 right). Fig. 1 (right) plots the Cumulative Distribution Functions of all seven blocks of the human data using the quantile boundaries from the first block. Note that as time on task increases (block number increases), the cumulative percentage of RTs falling below each boundary decreases, illustrating the shift of the distribution. The PDF and the CDF provide related information about the

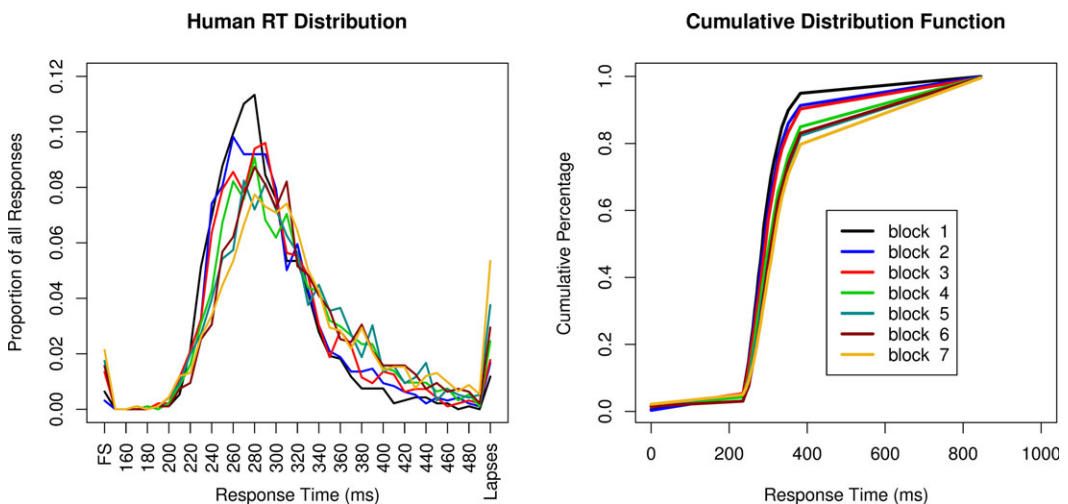


Fig. 1. Probability density functions (PDFs) and cumulative distribution functions (CDFs) for the PVT human response times from each 5-min block. In the CDF, Blocks 2–7 use the quantile boundaries from Block 1.

response time distribution. However, the CDF provides a concise description, is monotonic, and it captures the full distribution, which makes comparison across blocks and between model and human data much more straightforward.

5.4. Discussion

The 10-min PVT has been used extensively in the sleep research community to assess the impact that sleep deprivation has on cognitive performance. Within as little as one night of sleep deprivation, the percent of responses characterized as lapses and false starts increases significantly (e.g., Doran et al., 2001). To bridge the gap between the sleep and vigilance research communities, we extended the PVT to run for 35 min, a duration similar to tasks used in most vigilance studies. The results from the 35-min PVT demonstrate that performance degrades in a manner, and to a degree, similar to declines seen in the 10-min version of the PVT after a full night of sleep deprivation. In particular, we found that the increase in median RT over the 35-min session in our study was commensurate with the increase in median RT during a 10-min session for participants who had a full night of sleep deprivation (having been awake for 29 h) from a 62 h-sleep deprivation study (Tucker, Whitney, Belenky, Hinson, & Van Dongen, 2010). In the same study, the authors found that the percentage of lapses was about 1.5% at baseline across the 10 min and increased to 9.03% after a night of sleep deprivation, similar to our results of 1.17% at baseline and 5.23% at the end of the 35-min session. They also found the percentage of false starts at baseline was about 1.4% and rose to 2.6% after one night of sleep deprivation, which is similar to our results of 0.64% and 2.13%, respectively (Tucker et al., 2010).

This suggests that extended time on task may produce similar degradations in cognitive function to those seen with sleep loss. To further support this claim, the next section describes a model that performs the 35-min PVT using the same modeling framework and mechanisms as have previously been applied to the 10-min PVT in sleep-deprived participants (Gunzelmann, Gross et al., 2009; Walsh et al., 2014).

6. Model of the PVT

The PVT model⁵ is composed of three productions: *wait*, *attend*, and *respond* as illustrated in Fig. 2 and was run for 35 min. Initially, while no stimulus is present on the screen, the model executes the *wait* production (and potentially *respond*; see below). Once the stimulus appears, the model has to shift its attention to the stimulus source. This is achieved by executing the *attend* production, which sends a request to the visual module to shift attention to the stimulus. Once the attention shift has occurred, the *respond* production executes and initiates a *press key* action in the motor module.

Fig. 2 illustrates the general flow of the model. As there is noise introduced into cycle time and attention shift time, the approximate amount of time each cognitive action requires to complete is depicted. Note that the upper bound on the total time does not

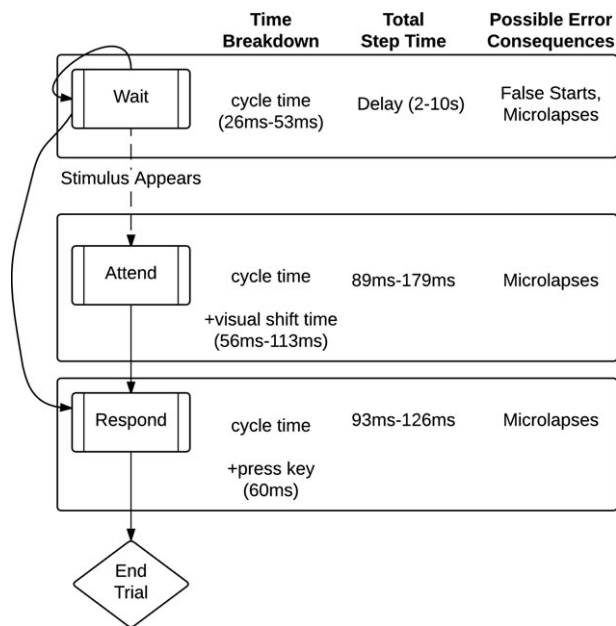


Fig. 2. Model of the Psychomotor Vigilance Test (PVT). Indicates the amount of time each production takes to execute and the possible implications of the fatigue module at each step.

include the impact of *microlapses*, which are instances where no productions reach threshold and therefore no cognitive actions are executed on a given cycle. *Microlapses* can occur during the conflict resolution phase in any step of model execution, and they become more prevalent as time on task increases.

False starts have the potential to occur while the model is waiting for the stimulus to appear. This is a direct result of the partial matching mechanism as the goal states for the *respond* and *wait* productions have a similarity = 0.5. Thus, although the *wait* production has an exact match to the goal state (and has a full utility), the *respond* production has a partial match with a discounted utility. We set the *ppm* parameter to twice the Initial Production Utility ($U_i(0)$) resulting in a mismatch penalty for the *respond* production during the waiting period calculated as follows: $U_i(t)' = IU - (2 \times IU) \times (0.5) + \epsilon = \epsilon$. *False starts* are thus a direct result of stochasticity introduced to the utility calculation. Because noise becomes more influential in the production selection and execution process with increased time on task, false starts become more likely.

6.1. Model evaluation

We explored the performance of the model by varying four parameters involved in calculating the utility and utility threshold, PU_{TOT} , UT_{TOT} , $UT(0)$, and $U_i(0)$, the ranges of which are shown in Table 2. We also used a non-default value for ACT-R's cognitive cycle time. This was done based on evidence that simple productions can execute more

Table 2
Parameter space run and resulting best fitting parameters for all human data in the 35-min Psychomotor Vigilance Test (PVT) model

Parameter	Range	Best Value
Production Utility Time on Task (PU_{TOT})	−0.01 to −1	−0.05
Utility Threshold Time on Task (UT_{TOT})	−0.01 to −1	−0.01
Initial Utility ($U_i(0)$)	1–5	2.6
Initial Utility Threshold ($UT(0)$)	1–5	2.2
Average CDF RMSE (R^2)		0.09 (0.98)
Average PDF RMSE (R^2)		0.01 (0.78)

quickly when implemented in a spiking neuron implementation of a production system architecture (Stewart, Choo, & Eliasmith, 2010). Their model predicted cycle times between 34 and 44 ms for “simple actions” like those instantiated by the productions in our model. This also aligns with previous modeling results with the PVT (Walsh et al., 2014), where a cycle time of 40 ms was found to best fit the human performance data on the standard 10-min version of the task. Because of this evidence, we adopted 40 ms as the cognitive cycle time for the model (the default value in ACT-R is 50 ms). Importantly, manipulations to this parameter do not provide a good account of performance changes on the PVT associated with sleep loss or circadian rhythms (see Gunzelmann, Byrne, Gluck, & Moore, 2009; Gunzelmann, Gross et al., 2009; Gunzelmann, Moore, et al., 2009, for a discussion). However, cognitive cycle time, interpreted as processing speed, has been shown to be useful in understanding individual differences in baseline performance on the PVT (Gunzelmann, Moore et al., 2009). Each parameter set was run 1,000 times in order to achieve stable performance estimates. The MindModeling@Home <http://mindmodeling.org/> resource was used in both the PVT and Mackworth model fitting (Harris, 2008).

The model’s performance was evaluated based on fit to the cumulative distribution functions (CDFs) for each 5-min block with the quantile boundaries based on the first block of the human data (see Fig. 1 right side). We compared human and model CDFs across all blocks and derived both root mean square error (RMSE) and R^2 values for each parameter set. The best fitting parameters are listed in Table 2 and the graphical representation of the changes in both Utility Threshold and Utility as a function of time-on-task is depicted in Fig. 3. The most pertinent aspect is the relationship between Utility and Utility Threshold—as time on task increases the difference between the two values decreases, leading to lower probabilities of successfully executing appropriate actions. As can be seen, both qualitatively in Fig. 4 and quantitatively in Table 2, the resulting fits to the extended PVT data were very good.

6.2. Traditional vigilance tasks versus PVT

We conducted an empirical study of a prolonged PVT to bridge the gap between the sleep deprivation and vigilance literatures. The degradations in performance in the prolonged PVT were similar to declines in performance observed after a night of sleep

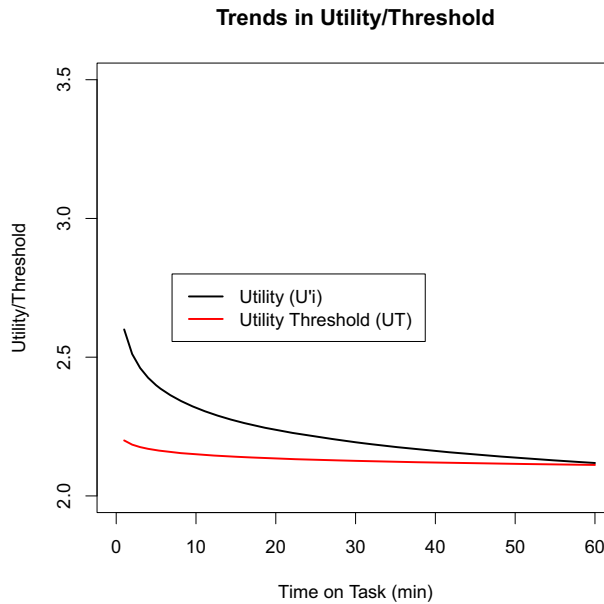


Fig. 3. Trends of production utilities and utility threshold of best fitting model parameters for Psychomotor Vigilance Test (PVT) model (prior to adding noise).

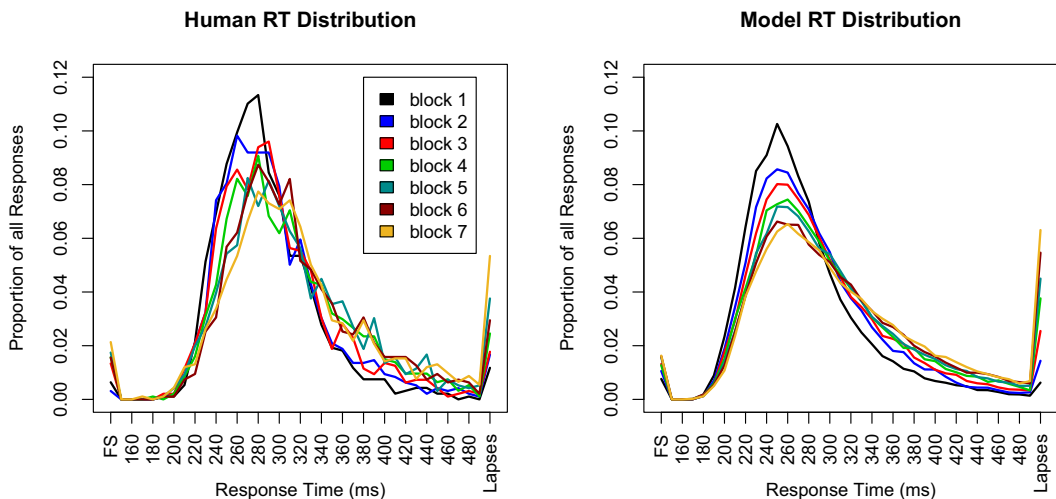


Fig. 4. Human and model RT distributions for all blocks.

deprivation, and they were captured by the same mechanisms originally described in the context of sleep loss. Furthermore, the degradation in performance appears similar to results from other vigilance tasks and studies.

These results are encouraging. However, while the 35-min PVT is more commensurate with typical vigilance tasks in its duration, there remain critical differences relative to

traditional vigilance tasks. First, participants in traditional vigilance studies must discriminate between signals and non-signals. In contrast, stimulus discrimination is not required in the PVT. Each time the counter appears on the screen, a response must be made.

Second, signals are rare in a traditional vigilance task, generally about 3% of stimuli. In the PVT, each stimulus presentation is a signal (100%), resulting in a high signal rate (around 80 signals per 10-min session). The low signal rate in traditional vigilance tasks is considered to be the primary factor influencing the decline in performance over time (only about 10–20 signals within a 30-min session). Generally speaking, lower signal rates lead to larger decrements (Davies & Parasuraman, 1982; Parasuraman, 1979).

Third, although signals are rare in traditional vigilance tasks, stimuli are typically presented at a rate of about 1 per second. Combined with a low signal rate, the high presentation rate of non-signals exacerbates the decrement (Lanzetta, Dember, Warm, & Berch, 1987; Wiener, 1977). On the other hand, during the PVT, stimuli are presented every 2–10 s, which is a much lower stimulus presentation rate.

Finally, response time is the only dependent measure provided by the PVT, as each stimulus requires a response. There is no measure of accuracy as defined by correctly identifying a signal, though lapses are sometimes considered errors in the task (e.g., Lim & Dinges, 2008). However, traditional vigilance tasks often report response time and accuracy measures, which may be differentially affected by time on task (Adams, Stenson, & Humes, 1961).

In summary, there are five significant differences between the PVT and traditional vigilance tasks: (a) whether stimulus discrimination is required; (b) signal probability; (c) stimulus rate; (d) dependent measures; and (e) task duration. Because our experiment addressed only one of those, it leaves questions unanswered about the capacity of the fatigue mechanism to account for the vigilance decrement. To bridge this gap, we turn next to a classic vigilance task, the Mackworth Clock Task (Mackworth, 1948).

7. Mackworth Clock Task

The Mackworth Clock Task (MCT) is the seminal task in the vigilance literature, run originally in the 1940s (Mackworth, 1948) and replicated in numerous studies (Giambra & Quilter, 1987; Kaida, Akerstedt, Kecklund, Nilsson, & Axelsson, 2007; Kass, Vodanovich, Stanny, & Taylor, 2001; Lichstein, Riedel, & Richman, 2000). Recent instantiations of the MCT have been used to generate datasets that are better suited for model comparison (Giambra & Quilter, 1987; Lichstein et al., 2000). For instance, Mackworth originally only reported average accuracy of 25 participants for each 30-min block in a 2-h session. Giambra and Quilter (1987), however, ran the study with 613 participants to obtain both accuracy and response time data for each of the 23 critical signals over the course of a 62-min study. We extended our model to perform the MCT and compared the model's performance to both the (limited) data of Mackworth's original study and the more robust data from Giambra and Quilter (1987).

7.1. Mackworth (1948)

Mackworth's (1948) original study of vigilance used a physical metal box with a plain white face and a single black pointer that moved in 100 discrete steps around the face (one tick per second). The double ticks occurred 12 times in 20 min followed by 10 min during which no signals occurred. The schedule of when the double jumps would occur was set at intervals of $\frac{3}{4}$, $\frac{3}{4}$, 1.5, 2, 2, 1, 5, 1, 1, 2, and 3 min. Participants in the 2-h study had the same schedule of four blocks of 30 min of the above schedule. Human data are presented in Fig. 6 (black open circles).

7.2. Giambra and Quilter (1987)

The replication done by Giambra and Quilter (1987) had participants complete 62 min of the Clock Task with 23 signals occurring during minutes 2, 4, 7, 11, 13, 15, 19, 21, 23, 26, 30, 32, 34, 37, 41, 43, 45, 49, 51, 53, 56, 60, and 62. Since all participants had the same signal schedule, data were aggregated over all 613 participants and presented in Fig. 8 (open circles). Giambra and Quilter (1987) reported that few subjects had false alarms, making mean values for that measure unreliable.

7.3. Model

Similar to the PVT model, the computational cognitive model for the Mackworth Clock Task is composed of three productions: *wait*, *attend*, and *respond*. The *wait* production executes between screen updates. The *attend* production executes when there is a screen change—when the clock hand moves to a new location around the clock. Visual attention is then directed to the location to which the clock hand moved by the model by executing the *attend* production. The *respond* production executes when the critical signal is detected—in this case, a hand movement of two ticks.

ACT-R's visual system does not actually sense the visual signal of the clock hand movement as the human eye would with the retina, so it was necessary to present a noisy, symbolic representation of the stimuli to ACT-R. To reflect the noise in perceiving the stimulus (the magnitude of the clock hand movement), we estimated sensitivity from the human data using signal detection theory (Green & Swets, 1966). While we realize that this was not a pure measure of visual sensitivity, it was the closest empirical measure possible in this context. To achieve this, the magnitude of clock hand movement of two ticks was drawn from a normal distribution with a mean of 2 and a standard deviation estimated from the human data at the beginning of the study. Neither Giambra and Quilter (1987) nor Mackworth (1948) reported sensitivity information for participants performing the MCT; so, we turned to Lichstein et al. (2000) who ran a computerized version of the Mackworth Clock Task and reported d' (sensitivity) for each half hour block, while also replicating Mackworth's accuracy results. Lichstein et al. (2000) indicated that in the first block of 30 min, the sensitivity value was 3.97, whereas on the

second block, it dropped down to 3.52. The relationship between sensitivity, d' , and the standard deviation, σ , of the signal distribution can be computed as:

$$d' = (\mu_s - \mu_n / \sigma) \tag{7}$$

where μ_s is the mean of the signal distribution (in this case, two ticks) and μ_n is the mean of the noise distribution (in this case, one tick). Therefore, in our case, $d' = 1/\sigma_n$.

The distribution of values used in the model had a mean equal to either 1 or 2 depending on the true clock hand movement and a $\sigma = 0.252$ around each of the “tick” values. This value for the parameter was found by calculating the d' obtained from the human data (Lichstein et al., 2000) in the first block of 30 min and applying it in Eq. 7 (see Fig. 5, left).

Production partial matching applies in this model in determining when to respond to a stimulus based upon the encoded tick magnitude (including noise). This allows the *respond* production to execute when the encoded magnitude is not exactly two ticks, which will generally be the case because of noisy encoding of the stimulus. The *ppm* parameter specifies the degree to which the production’s utility is discounted whenever the value is not exactly 2. As ACT-R does not provide a default value for *ppm*, the space of reasonable values was explored (in the range of 1–5, step size of 0.05). A *ppm* value of 2.5 produced the best results in anchoring the model to the first data point in the human data. The production partial matching mechanism requires a similarity calculation between the desired clock hand value (2) and the actual fuzzy signal in the visual buffer. This was defined as the negative of the absolute difference. Fig. 5 (right) depicts the utility value as a function of the encoded “tick” value with the *ppm* parameter set to 2.5 (the

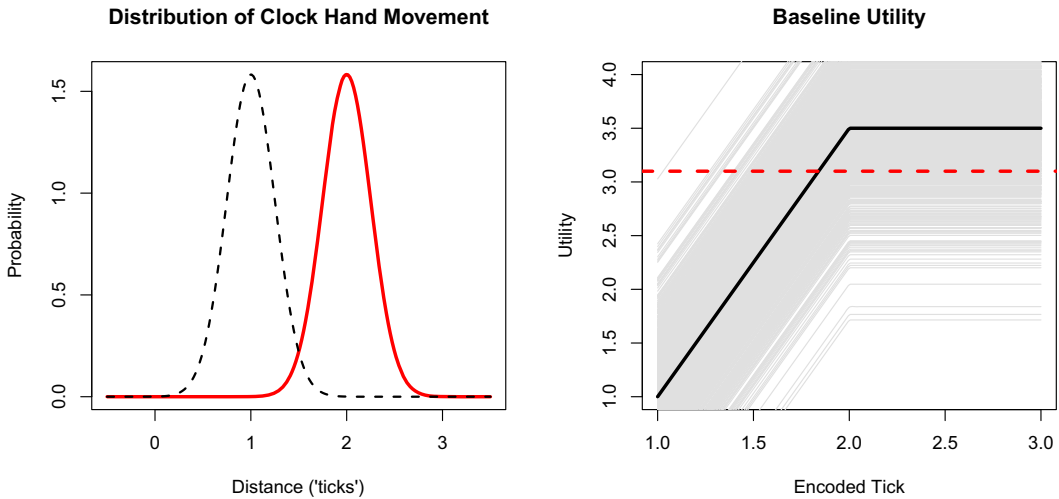


Fig. 5. Left panel: Distribution of values encoded by model around actual clock hand movement ($\sigma = 0.252$). Right panel: Utility value as a function of encoded “tick” value. Red dotted line indicates the initial utility threshold prior to any time-on-task effects.

noise-free value is indicated by the black line). The gray bands indicate possible range of utility values once ACT-R noise is added. The utility threshold of 3.1 (the best parameter setting for this data, as discussed below) is included for comparison.

There are two sources of error that could contribute to the model not responding correctly to the critical signal—the partially matched signal in the visual buffer and the inadvertent lack of a response because of microlapses. Since the model's sensory encoding accuracy remained constant during the entire duration of the task, the model's declining performance was a direct result of the microlapse mechanism affecting performance as time on task increased. The model embodies an important claim about the source of performance deficits associated with the vigilance decrement. Unlike many resource accounts of the vigilance decrement, the primary cause of degradations in behavior is not in the sensitivity of the perceptual and cognitive system to the target stimulus (see Thomson et al., 2015). Instead, it is disruptions in central cognitive control related to goal-directed processing that breaks down. Interestingly, this kind of deficit is more typical of *mind wandering* accounts of the vigilance decrement (Thomson et al., 2015).

7.4. Model evaluation

The model was given the same schedule of signals for each respective dataset. We first ran the parameter space comparing the model results with the original Mackworth accuracy measure (aggregated over 30-min blocks in a 2-h study). The parameters were systematically varied in precisely the same manner as in the PVT. The results are reported in Table 3. Each parameter set was run 3,000 times in order to increase confidence in the performance estimates. Fig. 6 plots human and model data for the original Mackworth Clock Task dataset (Mackworth, 1948).

Fig. 7 plots the relationship between Utility Threshold and Utility using the best parameter set (based on minimizing the RMSE between human and model data in accuracy). In the model, the difference between Initial Utility and Initial Utility Threshold is the primary driver of initial accuracy, while the relative difference between the PU_{TOT} and UT_{TOT} parameters is the main determinant of the magnitude of the vigilance decrement. A difference between Figs. 3 and 7 is that the Utility value falls below the Utility Threshold earlier in the Mackworth Clock Task model fitting. These differences are consistent with task differences, discussed above, that have been shown to influence the magnitude of the vigilance decrement. Specifically, both a higher stimulus presentation rate and a lower signal probability are factors that have been identified that exacerbate the vigilance decrement. Based on these factors, one would expect the MCT to produce a larger decrement than the PVT, and the best fitting parameters of the model suggest that this is the case. Of course, there are many other potential influences, including both individual and task factors, which are beyond the scope of this paper.

Having achieved a good fit to the original Mackworth data, we reran the model using the same parameters as above but compared the model's performance to the more detailed results reported in Giambra and Quilter (1987), which included 23 accuracy scores and 23 average RT values. This more detailed dataset allowed us to test the

Table 3
Parameter space explored and best fitting parameters for the 2-h Mackworth Clock Task model

Parameter	Range	Best Value
Production Utility Time on Task (PU_{TOT})	-0.01 to -1	-0.25
Utility Threshold Time on Task (UT_{TOT})	-0.01 to -1	-0.16
Initial Utility ($U_i(0)$)	1-5	3.6
Initial Utility Threshold ($UT(0)$)	1-5	3.1
Accuracy RMSE		0.01

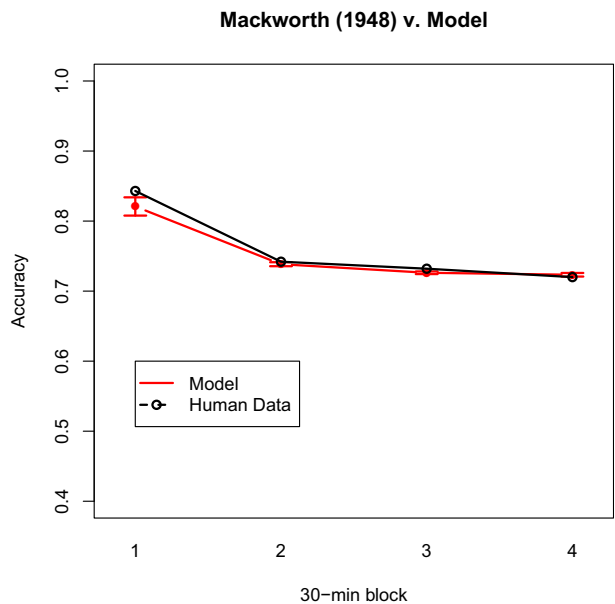


Fig. 6. Model fit to Mackworth (1948) data.

generalizability of the model in its ability to fit not only more fine grained accuracy data but response times as well. One caveat is that participants in Mackworth’s original study actually had better accuracy across the entire task than Giambra and Quilter’s participants. To compensate for this difference in overall magnitude, we adjusted the Utility Threshold by increasing it slightly to 3.2. This shifted all accuracy values down but did not alter the slope of the decline. An increased Utility Threshold resulted in fewer productions firing and led to more missed signals earlier in the task. Notionally, a lower threshold in the original Mackworth study may reflect higher motivation or effort on the part of the participants to perform well on the task.

Fig. 8 plots the results of the human and model data from the 62-min Giambra and Quilter (1987) study. In general, the model fit was good in terms of both accuracy (RMSE = 0.04; $R^2 = 0.95$) and response time (RMSE = 0.02; $R^2 = 0.71$). The model exhibited the same sharp initial decline in accuracy and then a leveling off as the human

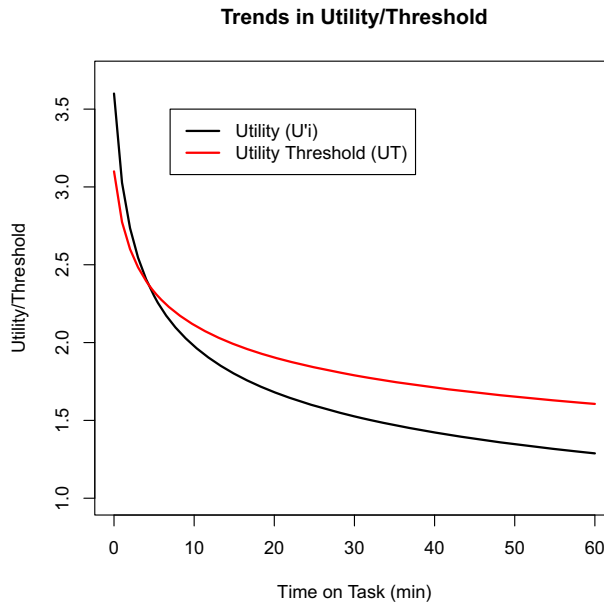


Fig. 7. Trends of production utilities and utility thresholds of best fitting model parameters for Mackworth (1948) model.

participants. As a further check of the mechanism, we extrapolated model performance to 2 h and observed that the model asymptotes at around 65% accuracy as in the human data (red triangle in Fig. 8).

8. General discussion

Computational cognitive models have made great strides in explaining and predicting performance on a variety of tasks (Anderson et al., 2004; Laird, 2012; Sun, 2006). However, these models typically disregard performance changes that may occur as a result of fluctuations in arousal and alertness due to sleep loss, time on task, and other cognitive moderators (i.e., caffeine, drugs, emotions). To make valid predictions about task performance, computational models need to have mechanisms that can capture how performance changes with respect to these factors (Gluck & Gunzelmann, 2013; Gluck, Gunzelmann, Gratch, Hudlicka, & Ritter, 2006). The introduction of mechanisms for fatigue within the ACT-R cognitive architecture is a good starting point for exploring the impact on central cognition and the consequences for task performance (Gunzelmann, Gross et al., 2009).

The research described here extends and validates fatigue mechanisms, originally developed to account for sleep loss, to the vigilance decrement, bridging the gap between sleep loss and vigilance fields of study and providing theoretical unification. The original fatigue mechanism in ACT-R modeled the effects of sleep loss and circadian rhythms on

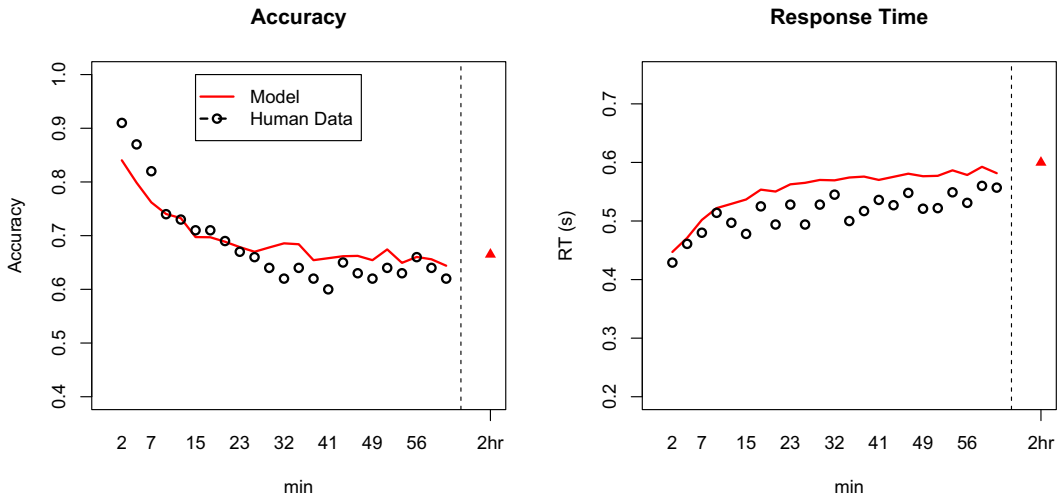


Fig. 8. Model fit to Giambra and Quilter (1987) data.

performance over the course of 88 h awake on a 10-min PVT (Gunzelmann, Gross et al., 2009). The mechanisms in that model instantiate a version of the state-instability hypothesis (Doran et al., 2001), producing changes in performance that have historically been described as separate phenomena of slowing and lapsing (e.g., Dinges & Kribbs, 1991).

We assessed the mechanism's generalizability by conducting a longer duration PVT in alert participants. Interestingly, we found that performance degraded over the course of 35 min in a manner similar to that found in participants after 24 h awake (Tucker et al., 2010). To truly test the robustness of the fatigue mechanism, we turned to the most traditional of vigilance tasks, the Mackworth Clock Task. Both accuracy and response time data in this task were captured by the model, in both cases due to the impact of the same mechanisms (namely *microlapses*) that accounted for sleep and circadian effects in prior research. This supports the conclusion that degradations in performance associated with the vigilance decrement are functionally equivalent to the declines associated with sleep loss.

From a theoretical standpoint, the mechanisms implement a resource account of the vigilance decrement. Resource depletion is reflected in reductions in utility values for productions in central cognition. Unlike most resource accounts, however, the consequences—*microlapses*—most directly impact task engagement rather than sensitivity to the stimulus. This kind of degradation is more consistent with mind wandering accounts of the vigilance decrement. However, whereas mind wandering is generally associated with understimulation and boredom (e.g., Damrad-Frye & Laird, 1989), our model demonstrates how similar consequences may emerge from the depletion of resources that are critical to cognitive control and goal-directed processing. This aligns with evidence suggesting that the vigilance decrement occurs despite increasing effort and higher workload in individuals performing those tasks (Warm et al., 2008).

This proposal is related, in some ways, to other recent attempts to integrate resource and mind wandering accounts of the vigilance decrement (Thomson et al., 2015). They use the term “executive control” to refer to the critical cognitive resource that is depleted in vigilance tasks, indicating that breakdowns in executive control lead to increased mind wandering. In our account, reductions in utility certainly lead to breakdowns in cognitive control. However, rather than producing mind wandering behavior, our model experiences microlapses in goal-directed processing. It would be possible to represent “mind wandering” in our model by implementing a larger and more diverse set of possible cognitive actions competing within ACT-R’s conflict resolution process. In fact, productions unrelated to the task of interest have been used in other contexts to account for inefficiencies in cognitive processing (e.g., Altmann & Gray, 2008).

It would be possible for both distracting thoughts and microlapses to interrupt cognitive processing within the same basic theoretical framework we have proposed. However, the addition of irrelevant knowledge and the associated complexity it introduces to the model is unnecessary in the current context to explain the empirical phenomena. Moreover, executing a task-irrelevant production is formally very similar to the occurrence of a microlapse, and so the mechanisms and the explanation would remain largely consistent. Specifically time on task, like sleep loss, decreases utility (proportionally), which reduces the signal-to-noise ratio in the selection and execution of cognitive actions. Because our mechanisms span vigilance and sleep loss, the resolution could depend on theoretical arguments regarding the nature of cognitive processing during sleep (e.g., Diekelmann & Born, 2010; Siegel, 2005). Thus, we leave the debate to future research.

Distinguishing between distraction and microlapses is an instance of the more general challenge of identifiability. Many alternative mechanisms could be considered to account for the kinds of declines observed with sleep loss, circadian rhythms, and time on task. Gunzelmann, Gross et al. (2009) explored various other potential mechanisms that could have accounted for changes in cognitive function due to sleep deprivation. Namely, they explored increased cognitive noise, cognitive slowing, and a combination of the two. None of these alternatives provided a good account of the empirical data, however. Though some of these mechanisms could account for some of the effects observed in the Mackworth Clock Test, we have adopted the parsimonious approach of generalizing previously validated mechanisms to new, related phenomena, rather than proposing new mechanisms.

The ability to generalize the mechanisms to new tasks is at the core of the motivation for using a cognitive architecture. It may be possible to account generally for the response time distributions in the PVT with mathematical functions (e.g., a gamma distribution). However, such accounts are limited to “descriptive curve fitting” (Dinges, 2004, p. A182), with no ability to make principled predictions beyond the tasks where the parameters have been fit. For instance, in the current research, fitting functions to the PVT distribution provides no insight to the decline in the MCT. This lack of generalizability undermines a significant opportunity with research in this area, which is making predictions in contexts where fatigue is a significant risk to safety and data are not available to drive parameter fitting or conclusions about how to mitigate that risk (Gunzelmann et al., 2015).

For a large portion of the history of cognitive modeling, the goal has been to have a parsimonious account of the ideal performer or expert. Although this is a worthy goal, a complete theory of human cognition must go further. In Newell's (1990) terms, our research moves theories of cognition "further down the list cumulatively than we have ever done before" (p. 16), in accounting for both the capacities and limitations of the human mind. By being able to capture the fluctuations in performance that often impact people outside the laboratory, we can paint a much more veridical picture of the human mind. In addition, understanding these phenomena also creates opportunities to apply the research to mitigate the risks associated with the deleterious effects of fatigue (e.g., Gunzelmann et al., 2015). It is the case that sleep loss leads to fatigue, which in turn leads to decrements in performance. At the same time, prolonged time spent performing monitoring tasks leads to fatigue as well, which is exacerbated by sleep loss (e.g., Doran et al., 2001; McIntire, McKinley, Goodyear, & Nelson, 2014). In order to produce parsimonious models that capture various aspects of performance, it would be beneficial to have a unified mechanism that can account for both effects seen at the longer time scale of hours or days, caused by sleep restriction or deprivation, and changes that unfold at a shorter scale of minutes to hours, as seen in vigilance studies.

Our work demonstrates how the progressive declines in utility in ACT-R's central production system can lead to performance changes consistent with observed effects in human behavior related to fatigue in its various forms. The degradations can be seen as resulting from the overuse of cognitive processing capacities, particularly top-down attentional control. In the model, this leads to decrements in utility, which represent the primary resource in ACT-R related to the selection and execution of cognitive actions. Depleted utility in central cognition leads to microlapses in goal-directed processing. Microlapses, in turn, are responsible for performance delays (i.e., *cognitive slowing*; Bratzke, Rolke, Ulrich, & Peters, 2007), and more severe disruptions (i.e., *cognitive lapses*; Kjellberg, 1977). The compensatory mechanism—reduced utility threshold—mitigates these effects, but also increases the likelihood of inappropriate actions, resulting in false alarms in the data presented here. This provides an explanation for why struggling to maintain performance while fatigued can lead to breakdowns in inhibition, producing greater variability in behavior and performance (e.g., Chuah et al., 2006).

In addition to supporting current theories of the vigilance decrement, the model is also consistent with research showing that the responsiveness of neural populations declines as neurons "tire out" (Krueger et al., 2008). In both cases, the likelihood of a response is diminished. As neuron populations "tire out," the system gravitates toward non-responsiveness. Some evidence suggests that the neural structures associated with sleep and circadian rhythms function to coordinate neural rest and recovery, so that performance can be maintained better during wakefulness (Saper et al., 2001). This further supports the equivalence of fatigue brought on by sleep loss (preventing recovery) and time on task (increasing recovery need through overuse). The current research identifies a candidate resource and describes a mechanism to account for the impact that depletion of that resource has on information processing. Additional research is needed to better understand the mechanisms of recovery, how cognitive resources are managed during more

typical cognitive activities outside the laboratory, and how other factors like expertise and motivation influence these dynamics.

The interplay of these issues is apparent in the model, where the parameters used for the PVT and Mackworth models are different. One reason for this is that the task contexts are different, which may be associated with task difficulty, motivation, and other factors. Additionally even in the context of the Mackworth task, baseline performance of participants in Mackworth's (1948) study was noticeably better than Giambra and Quilter's (1987), despite similar methodologies. Further validation of the fatigue mechanisms in other vigilance task contexts is needed, as is a better understanding of the interaction of individual differences, the task context, and the environment with alertness and cognitive performance to understand these important, but complicated, dynamics.

In conclusion, by validating that mechanisms for sleep loss can also account for the negative consequences of time on task, we have demonstrated the functional equivalence of the declines in performance observed in both domains, at least for sustained attention tasks. The mechanisms themselves are consistent with current psychological theory and neuroscientific evidence in both domains, and they account well for the empirical phenomena. The research aids in unifying theories regarding these disparate sources of fatigue and contributes to the development of more comprehensive theories of human cognition more generally. It helps to inform a closer connection between the domains of sleep deprivation and vigilance and illustrates how computational cognitive models can incorporate quantitative theories of how cognitive moderators influence performance, to better account for the nuances of human cognition and mind.

Acknowledgments

The views in this paper reflect those of the authors, and are not the official views of the United States Air Force. We thank Monica Nguyen for supporting data collection; Jack Harris for experiment setup and MindModeling@Home support; and Kevin Gluck, Matthew Walsh, and Rachel Boyd for providing insightful comments on earlier drafts of this manuscript.

This work was performed while the first author held a National Research Council Research Associateship Award at Air Force Research Lab. This work was supported by the Air Force Office of Scientific Research (AFOSR) (grant number 09RH06COR).

Notes

1. Models and their corresponding publications can be found at <http://act-r.psy.cmu.edu/publication/>
2. This utility calculation has been modified to maintain currency with the evolution of ACT-R. The current model code is written and tested in ACT-R version 6.0, released December 2014.

3. Estimating parameters associated with sleep and circadian rhythm requires multiple observations taken at different points in time. In the model, these parameters would affect the starting value of FP , which affect the $U_i(t)$ value prior to the impact of time on task. Because these parameters cannot be estimated in the current experiment, we estimate starting values for $U_i(t)$ directly. In this model, $U_i(t)$ should be thought of as representing $FP'(0) \times U_i(0)$, where FP' is the FP value with sleep and circadian rhythms included, and $U_i(0)$ is the true non-fatigued utility value.
4. Sleep debt accumulates over the course of multiple nights and effects of caffeine depend on various factors such as timing, dose, and tolerance (Ramakrishnan et al., 2014; Van Dongen, Baynard, Maislin, & Dinges, 2004; Van Dongen et al., 2001). While both sleep and caffeine can impact performance due to their effect on alertness, in this study these factors were not precisely measured or controlled. Thus, it is not particularly surprising that no effects were found.
5. The model described here varies in some respects from the model in Gunzelmann, Gross et al. (2009). Changes were made to maintain currency with the evolution of the ACT-R architecture. The mechanisms remain theoretically consistent and produce equivalent fits to human data (Walsh et al., 2014). The model and fatigue module code will be available as a supplement to this manuscript; both are intended for use with ACT-R 6.0, December 2014 release.

References

- Adams, J. A., & Boulter, L. R. (1960). *Monitoring of complex visual displays: I. Effects of response complexity and Intersignal Interval on Vigilant Behavior when Visual Load is Moderate*. Urbana, IL: University of Illinois.
- Adams, J. A., & Boulter, L. R. (1964). Spatial + temporal uncertainty as determinants of vigilance behavior. *Journal of Experimental Psychology*, 67(2), 127–131.
- Adams, J. A., Stenson, H. H., & Humes, J. M. (1961). Monitoring of complex visual displays—II. effects of visual load and response complexity on human vigilance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 3(4), 213–221.
- Altmann, E. M., & Gray, W. D. (2008). An integrated model of cognitive control in task switching. *Psychological Review*, 115(3), 602–639.
- Anderson, J. R. (2007). *How can the human mind exist in the physical universe?* New York: Oxford University Press.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, 111(4), 1036–1060.
- Bakan, P. (1955). Discrimination decrement as a function of time in a prolonged vigil. *Journal of Experimental Psychology*, 50(6), 387.
- Basner, M., & Dinges, D. F. (2011). Maximizing sensitivity of the psychomotor vigilance test (PVT) to sleep loss. *Sleep*, 34(5), 581–591.
- Belavkin, R. V., Ritter, F., & Elliman, D. G. (1999). Towards including simple emotions in a cognitive architecture in order to better fit children's behavior. In M. Hahn and Scott C. (Eds.), *Proceedings of the 1999 Conference of the Cognitive Science Society* (p. 784). Mahwah, NJ: Lawrence Erlbaum.

- Benitez, P. L., Kamimori, G. H., Balkin, T. J., Greene, A., & Johnson, M. L. (2009). Modeling fatigue over sleep deprivation, circadian rhythm, and caffeine with a minimal performance inhibitor model. *Methods in Enzymology*, 454 (C), 405–421.
- Bevan, W., Avant, L. L., & Lankford, H. G. (1967). Influence of interpolated periods of activity and inactivity upon the vigilance decrement. *Journal of Applied Psychology*, 51(4), 352–356.
- Bratzke, D., Rolke, B., Ulrich, R., & Peters, M. (2007). Central slowing during the night. *Psychological Science*, 18(5), 456–461.
- Chuah, Y. M. L., Venkatraman, V., Dinges, D. F., & Chee, M. W. L. (2006). The neural basis of interindividual variability in inhibitory efficiency after sleep deprivation. *Journal of Neuroscience*, 26(27), 7156–7162.
- Damrad-Frye, R., & Laird, J. D. (1989). The experience of boredom: The role of the self-perception of attention. *Journal of Personality and Social Psychology*, 57(2), 315–320.
- Davies, D., & Parasuraman, R. (1982). *The psychology of vigilance*. London: Academic Press.
- Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, 11(2), 114–126.
- Dinges, D. F. (2004). Critical research issues in development of biomathematical models of fatigue and performance. *Aviation Space and Environmental Medicine*, 75(3), A181–A191.
- Dinges, D. F., & Kribbs, N. B. (1991). Performing while sleepy: effects of experimentally-induced sleepiness. In T. H. Monk (Ed.), *Sleep sleepiness and performance* (pp. 98–128). Chichester, England: John Wiley & Sons.
- 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.
- Doran, S. M., Van Dongen, H. P. A., & Dinges, D. F. (2001). Sustained attention performance during sleep deprivation: Evidence of state instability. *Archives of Italian Biology: Neuroscience*, 139(3), 253–267.
- Drummond, S. P., Brown, G. G., Gillin, J. C., Stricker, J. L., Wong, E. C., & Buxton, R. B. (2000). Altered brain response to verbal learning following sleep deprivation. *Nature*, 403, 655–657.
- Drummond, S. P., Paulus, M. P., & Tapert, S. F. (2006). Effects of two nights sleep deprivation and two nights recovery sleep on response inhibition. *Journal of Sleep Research*, 15(3), 261–265.
- Duffy, J. F., Kronauer, R. E., & Czeisler, C. A. (1996). Phase-shifting human circadian rhythms: Influence of sleep timing, social contact and light exposure. *The Journal of Physiology*, 495(Pt 1), 289–297.
- Fisk, A. A. D., & Schneider, W. (1981). Control and automatic processing during tasks requiring sustained attention: A new approach to vigilance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 23(6), 737–750.
- Gartenberg, D., Veksler, B. Z., Gunzelmann, G., & Trafton, J. G. (2014). An ACT-R process model of the signal duration phenomenon of vigilance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 58(1), 909–913.
- Giambra, L. M., & Quilter, R. E. (1987). A two-term exponential functional description of the time course of sustained attention. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 29(6), 635–643.
- Gluck, K. A., & Gunzelmann, G. (2013). Computational process modeling and cognitive stressors: background and prospects for application in cognitive engineering. In J. D. Lee & A. Kirlik (Eds.), *The Oxford handbook of cognitive engineering* (pp. 424–432). New York: Oxford University Press.
- Gluck, K. A., Gunzelmann, G., Gratch, J., Hudlicka, E., & Ritter, F. (2006). Modeling the impact of cognitive moderators on human cognition and performance. In R. Sun & N. Miyake (Eds.), *Twenty-eighth annual meeting of the cognitive science society* (pp. 26–58), Mahwah, NJ: Lawrence Earlbaum Associates.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Gunzelmann, G., Byrne, M. D., Gluck, K. A., & Moore, L. R. (2009). Using computational cognitive modeling to predict dual-task performance with sleep deprivation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 51(2), 251–260.

- Gunzelmann, G., Gross, J. B., Gluck, K. A., & Dinges, D. F. (2009). Sleep deprivation and sustained attention performance: Integrating mathematical and cognitive modeling. *Cognitive Science*, 33(5), 880–910.
- Gunzelmann, G., Moore, L. R., Gluck, K. A., Van Dongen, H. P. A., & Dinges, D. F. (2009). Examining sources of individual variation in sustained attention. In N. Taatgen & van Rijn H. (Eds.), *Thirty-first annual meeting of the cognitive science society* (pp. 608–613). Austin, TX: Cognitive Science Society.
- Gunzelmann, G., Moore, L. R., Gluck, K. A., Van Dongen, H. P. A., & Dinges, D. F. (2010). Fatigue in sustained attention: Generalizing mechanisms for time awake to time on task. In P. L. Ackerman (Ed.), *Cognitive fatigue: Multidisciplinary perspectives on current research and future applications* (pp. 83–96). Washington, DC: American Psychological Association.
- Gunzelmann, G., Moore, L. R., Salvucci, D. D., & Gluck, K. A. (2011). Sleep loss and driver performance: Quantitative predictions with zero free parameters. *Cognitive Systems Research*, 12(2), 154–163.
- Gunzelmann, G., Veksler, B. Z., Walsh, M. M., & Gluck, K. A. (2015). Understanding and predicting the cognitive effects of sleep loss through simulation. *Translational Issues in Psychological Science*, 1(1), 106–115.
- Harris, J. (2008). MindModeling@Home: A large-scale computational cognitive modeling infrastructure. In *Proceedings of the Sixth Annual Conference on Systems Engineering Research 2008* (pp. 246–252). Los Angeles, CA: The University of Southern California.
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, 52(C), 139–183.
- Hitchcock, E. M., Dember, W. N., Warm, J. S., Moroney, B. W., See, J. E., & Services, L. T. (1999). Effects of cueing and knowledge of results on workload and boredom in sustained attention. *Human Factors*, 41(3), 365–372.
- Hitchcock, E. M., Warm, J. S., Matthews, G., Dember, W. N., Shear, P. K., Tripp, L. D., Mayleben, D. W., & Parasuraman, R. (2003). Automation cueing modulates cerebral blood flow and vigilance in a simulated air traffic control task. *Theoretical Issues in Ergonomics Science*, 4(1), 89.
- Jerison, H. J., & Pickett, R. M. R. (1964). Vigilance : The importance of the elicited observing rate. *Science*, 143(3609), 970–971.
- Jewett, M. E., & Kronauer, R. E. (1999). Interactive mathematical models of subjective alertness and cognitive throughput in humans. *Journal of Biological Rhythms*, 14(6), 588–597.
- Kaida, K., Akerstedt, T., Kecklund, G., Nilsson, J. P., & Axelsson, J. (2007). Use of subjective and physiological indicators of sleepiness to predict performance during a vigilance task. *Industrial Health*, 45(4), 520–526.
- Kass, S. J., Vodanovich, S. J., Stanny, C. J., & Taylor, T. M. (2001). Watching the clock: Boredom and vigilance performance. *Perceptual and Motor Skills*, 92(3), 969–976.
- Kennedy, R., & Coulter, X. (1975). Research note: The interactions among stress, vigilance, and task complexity. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 17(1), 106–109.
- Kilpeläinen, A. A., Huttunen, K. H., Lohi, J. J., & Lyytinen, H. (2010). Effect of caffeine on vigilance and cognitive performance during extended wakefulness. *The International Journal of Aviation Psychology*, 20(2), 144–159.
- Kjellberg, A. (1977). Sleep deprivation and some aspects of performance: II, Lapses and other attentional effects. *Waking and Sleeping*, 1, 145–148.
- Kleitman, N. (1963). *Sleep and Wakefulness*. Chicago: The University of Chicago Press.
- Krueger, J. M., Rector, D. M., Roy, S., Van Dongen, H. P. A., Belenky, G., & Panksepp, J. (2008). Sleep as a fundamental property of neuronal assemblies. *Nature Reviews Neuroscience*, 9(12), 910–919.
- Laird, J. E. (2012). *The Soar Cognitive Architecture*. Cambridge, MA: MIT Press.
- Lanzetta, T. M. T., Dember, W. N., Warm, J. S., & Berch, D. B. (1987). Effects of task type and stimulus heterogeneity on the event rate function in sustained attention. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 29(6), 625–633.
- Leproult, R., Colicchia, E. F., Berardi, A. M., Stickgold, R., Kosslyn, S. M., & Van Cauter, E. (2003). Individual differences in subjective and objective alertness during sleep deprivation are stable and

- unrelated. *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology*, 284, 280–290.
- Levine, J. (1966). The effects of values and costs on the detection and identification of signals in auditory vigilance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 8(6), 525–537.
- Lichstein, K. L., Riedel, B. W., & Richman, S. L. (2000). The mackworth clock test: A computerized version. *The Journal of Psychology*, 134(2), 153–161.
- Lim, J., & Dinges, D. F. (2008). Sleep deprivation and vigilant attention. *Annals of The New York Academy Of Sciences*, 1129(1), 305–322.
- Lim, J., & Dinges, D. F. (2010). A meta-analysis of the impact of short-term sleep deprivation on cognitive variables. *Psychological Bulletin*, 136(3), 375–389.
- Loeb, M., & Binford, J. R. (1964). Vigilance for auditory intensity changes as a function of preliminary feedback and confidence level. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 6(5), 445–458.
- Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology*, 1, 6–21.
- Mackworth, J. F. (1964). Performance decrement in vigilance, threshold, and high-speed perceptual motor tasks. *Canadian Journal of Psychology*, 18(3), 209–223.
- Mallis, M. M., Mejdal, S., Nguyen, T. T., & Dinges, D. F. (2004). Summary of the key features of seven biomathematical models of human fatigue and performance. *Aviation Space and Environmental Medicine*, 75(3 Suppl), A4–A14.
- McCauley, P., Kalachev, L. V., Mollicone, D. J., Banks, S., Dinges, D. F., & Van Dongen, H. P. (2013). Dynamic circadian modulation in a biomathematical model for the effects of sleep and sleep loss on waking neurobehavioral performance. *Sleep* 36(12), 1987–1997.
- McIntire, L. K., McKinley, R. A., Goodyear, C., & Nelson, J. (2014). A comparison of the effects of transcranial direct current stimulation and caffeine on vigilance and cognitive performance during extended wakefulness. *Brain Stimulation*, 7(4), 499–507.
- McVay, J. C., & Kane, M. J. (2010). Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). *Psychological Bulletin*, 136(2), 188–197; discussion 198–207.
- Molloy, R., & Parasuraman, R. (1996). Monitoring an automated system for a single failure: Vigilance and task complexity effects. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 38(2), 311–322.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
- Parasuraman, R. (1979). Memory load and event rate control sensitivity decrements in sustained attention. *Science*, 205(4409), 924–927.
- Portas, C. M., Rees, G., Howseman, A. M., Josephs, O., Turner, R., & Frith, C. D. (1998). A specific role for the thalamus in mediating the interaction of attention and arousal in humans. *Journal of Neuroscience*, 18(21), 8979–8989.
- Ramakrishnan, S., Laxminarayan, S., Wesensten, N. J., Kamimori, G. H., Balkin, T. J., & Reifman, J. (2014). Dose-dependent model of caffeine effects on human vigilance during total sleep deprivation. *Journal of Theoretical Biology*, 358, 11–24.
- Ramakrishnan, S., Rajaraman, S., Laxminarayan, S., Wesensten, N. J., Kamimori, G. H., Balkin, T. J., & Reifman, J. (2013). A biomathematical model of the restoring effects of caffeine on cognitive performance during sleep deprivation. *Journal of Theoretical Biology*, 319, 23–33.
- Robinson, E. S., & Bills, A. G. (1926). Two factors in the work decrement. *Journal of Experimental Psychology*, 9(6), 415–443.
- Saper, C. B., Chou, T. C., & Scammell, T. E. (2001). The sleep switch: Hypothalamic control of sleep and wakefulness. *Trends in Neurosciences*, 24(12), 726–731.
- Siegel, J. M. (2005). Clues to the functions of mammalian sleep. *Nature*, 437(7063), 1264–1271.

- Smith, R. L., Lucaccini, L. F., & Epstein, M. H. (1967). Effects of monetary rewards and punishments on vigilance performance. *Journal of Applied Psychology*, 51(5), 411.
- Smith, R. L., Lucaccini, L. F., Groth, H., & Lyman, J. (1966). Effects of anticipatory alerting signals and a compatible secondary task on vigilance performance. *Journal of Applied Psychology*, 50(3), 240.
- Stewart, T. C., Choo, X., & Eliasmith, C. (2010a). Dynamic behaviour of a spiking model of action selection in the basal ganglia. In D. Salvucci, & G. Gunzelmann (Eds.), *Proceedings of the 10th International Conference on Cognitive Modeling* (pp. 235–240). Philadelphia, PA: Drexel University.
- Sun, R. (2006). The CLARION cognitive architecture: Extending cognitive modeling to social simulation. In: R. Sun (Ed.). *Cognition and MultiAgent Interaction* (pp. 79–99). New York: Cambridge University Press.
- Teichner, W. H. (1974). The detection of a simple visual signal as a function of time of watch. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 16(4), 339–352.
- Temple, J. G., Warm, J. S., Dember, W. N., Jones, K. S., LaGrange, C. M., & Matthews, G. (2000). The effects of signal salience and caffeine on performance, workload, and stress in an abbreviated vigilance task. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 42(2), 183–194.
- Thomson, D. R., Besner, D., & Smilek, D. (2015). A resource-control account of sustained attention: Evidence from mind-wandering and vigilance paradigms. *Psychological Review*, 10(1), 82–96.
- Tucker, A. M., Whitney, P., Belenky, G., Hinson, J. M., & Van Dongen, H. P. a (2010). Effects of sleep deprivation on dissociated components of executive functioning. *Sleep*, 33(1), 47–57.
- Van Dongen, H. P. A. (2004). Comparison of mathematical model predictions to experimental data of fatigue and performance. *Aviation, Space, and Environmental Medicine*, 75(3 Suppl), A15–A36.
- Van Dongen, H. P. A., Baynard, M. D., Maislin, G., & Dinges, D. F. (2004). Systematic interindividual differences in neurobehavioral impairment from sleep loss: Evidence of trait-like differential vulnerability. *Sleep* 27(3), 423–433.
- Van Dongen, H. P. A., Belenky, G., & Krueger, J. M. (2011). A local, bottom-up perspective on sleep deprivation and neuro- behavioral performance. *Sleep Rochester*, 11(19), 2414–2422.
- Van Dongen, H. P. A., & Dinges, D. F. (2005). Sleep, circadian rhythms, and psychomotor vigilance. *Clinics in Sports Medicine*, 24(2), 237–249.
- Van Dongen, H. P. A., & Dinges, D. F. (2005b). Sleep, circadian rhythms, and psychomotor vigilance. *Clinics in Sports Medicine* 24(2). pp. 237–249, vii–viii.
- Van Dongen, H. P. A., Price, N. J., Mullington, J. M., Szuba, M. P., Kapoor, S. C., & Dinges, D. F. (2001). Caffeine eliminates psychomotor vigilance deficits from sleep inertia. *Sleep*, 24(7), 813–819.
- Van Zandt, T. (2000). How to fit a response time distribution. *Psychonomic Bulletin & Review*, 7(3), 424–465.
- Walsh, M. M., Gunzelmann, G., & Van Dongen, H. P. A. (2014). Comparing accounts of psychomotor vigilance impairment due to sleep loss. In P. Bello, M. Guarini, M. McShane & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 877–882). Austin, TX: Cognitive Science Society.
- Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(3), 433–441.
- Wiener, E. L. (1977). Stimulus presentation rate in vigilance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 19(3), 301–303.
- Wiener, E. L., Curry, R. E., & Faustina, M. L. (1984). Vigilance and task load: In search of the inverted U. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 26(2), 215–222.
- Williges, R. C. (1973). Manipulating the response criterion in visual monitoring. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 15(2), 179–185.