A Neurobehaviorally Inspired ACT-R Model of Sleep Deprivation: Decreased Performance in Psychomotor Vigilance

Glenn Gunzelmann (glenn.gunzelmann@mesa.afmc.af.mil)
Kevin A. Gluck (kevin.gluck@mesa.afmc.af.mil)
Air Force Research Laboratory; 6030 South Kent Street; Mesa, AZ 85212 USA

Hans P. A. Van Dongen (vdongen@mail.med.upenn.edu)
Robert M. O'Connor (roconnor@mail.med.upenn.edu)
David F. Dinges (dinges@mail.med.upenn.edu)
Division of Sleep and Chronobiology, University of Pennsylvania School of Medicine
1013 Blockley Hall; 423 Guardian Drive; Philadelphia, PA 19104-6021

Abstract

This paper describes how changes in architectural parameters in ACT-R can be used to understand and predict the effects of sleep deprivation on a fundamental aspect of human performance. In a sample task, the parameter manipulations produce changes in the model's performance that closely resemble the neurobehavioral effects seen in human data. The parameter that is manipulated (*G*) influences a mechanism in ACT-R that is considered to be associated with the thalamus, an area that is sensitive to sleep deprivation.

Introduction

Sleep is essential for normal human functioning. When people are deprived of sleep or even experience restricted sleep schedules, their performance degrades. These performance drops are evident in everything from simple sustained attention reaction time tasks like the psychomotor vigilance task (Van Dongen, 2004) to complex, dynamic tasks like flying high-fidelity military aircraft simulators (Caldwell et al., 2004). The effects of sleep deprivation range from subtle increases in reaction times to "sleep attacks," where an individual falls asleep while engaged in goal-directed behavior (Durmer & Dinges, 2005). These effects can have major consequences in settings where swing shifts or long or unusual hours are the norm, such as long-haul trucking, commercial aviation, and military operations.

In order to improve our ability to predict how and when performance will decline as a result of restricted or deprived sleep, we must improve our understanding of how sleep deprivation impacts the cognitive system. If these predictions can be made, then actions can be taken *a priori* to mitigate the effects of sleep deprivation and minimize the likelihood of costly or tragic fatigue-related errors (Dinges, 2004).

This paper describes our recent efforts along these lines. We first describe relevant research from the sleep restriction and cognitive modeling literatures. This is followed by a description of a cognitive model that makes predictions in a sample task used frequently in sleep restriction research.

The model is inspired by neurophysiological findings from the sleep restriction community, and demonstrates the promise of this approach for understanding the impact of sleep restriction on performance.

Neurophysiology of Sleep Restriction

Research on sleep restriction has identified some of the ways in which lack of sleep impacts brain activity (e.g., Durmer & Dinges, 2005; Drummond & Brown, 2001; Lin, 2000; Portas et al., 1998). The effects that are observed depend on what task is being performed, especially at the cortical level (e.g., Drummond & Brown, 2001). At the subcortical level, the thalamus has been implicated in regulating arousal (Lin, 2000; Moruzzi & Magoun, 1949), particularly in modulating attention (Portas et al., 1998). Research on sleep restriction has found changes in the activation of the thalamus as a function of sleep debt (e.g., Portas et al., 1998). These findings point to the thalamus as a key neural structure in mediating the effects of sleep deprivation on cognitive performance.

Math Models of Cognitive Throughput

Research on the effects of sleep restriction also has resulted in a better understanding of how sleep and circadian rhythms interact to influence an individual's ability to perform tasks. These findings have been incorporated into a variety of biomathematical models that are either commercially or publicly available (Mallis et al., 2004). A review of these models, their implementations, and their ability to predict novel empirical results was recently conducted and published (*Aviation, Space, and Environmental Medicine, 75,* 3, March 2004). All of the models produce some form of prediction of sleepiness or impairment in cognitive performance due to sleep loss, which is useful for quantifying the overall effectiveness of a person's cognitive system relative to maximal. However, these models do not make performance predictions in specific task situations. For example a model may predict only that cognitive throughput is at 70%, leaving it unclear

what the implications are for changes in response times or error rates on a particular task.

Computational Models of Cognition

Cognitive architectures like ACT-R and Soar provide an alternative to mathematical models in the cognitive scientist's toolbox. Architecture-based models allow the modeler to simulate cognitive processes in the context of a specific task domain. Owing to the "embodiment" movement (primarily the addition of visual and motor system representations) in the 1990's, modelers can now develop process models that interact with the same computer-based tasks or simulations used by participants in empirical studies, which greatly facilitates model validation.

A recent development is that researchers have started using these architectures as tools for explaining the effects of cognitive moderators on performance. For instance, the ACT-R architecture was used by Jongman (1998) to simulate within-task mental fatigue; by Belavkin (2001) to simulate the role of emotion in decision making; and by Ritter and colleagues (2004) to simulate pre-task appraisal and anxiety. These efforts reflect an increasing maturity in the architecture and a vibrant user community.

Another recent development within the ACT-R architecture has been an effort to establish a mapping between mechanisms in the architecture and regions in the brain. The current conceptualization of this mapping is shown in Figure 1. Research using fMRI has validated this mapping by demonstrating that increased activity in the brain regions of human subjects performing a task corresponds to increased activity in those components of a model performing the same task (Anderson et al., 2004).

Modeling Goal The major goal of our research is to produce a model that provides a principled computational account of how loss of sleep impacts the cognitive system. This paper reports recent successes at modeling the effects of sleep restriction in a sample task using ACT-R. We manipulate The *G* parameter, which is associated with production selection and execution in ACT-R (Figure 1), which aligns the modeling work with the findings from the sleep restriction community by focusing on a mechanism in ACT-R that is associated with the thalamus. The model's performance mirrors aspects of the impact of sleep deprivation in human participants. In the remainder of the paper, the task used in this research is described, followed by a detailed description of the model. The paper ends with conclusions and some directions for future research.
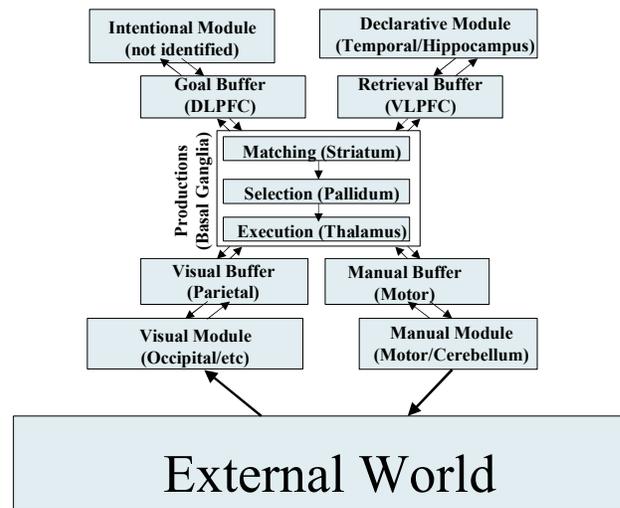

Figure 1: ACT-R Architecture, including mapping of ACT-R components to brain regions. Adapted from http://act-r.psy.cmu.edu/tutorials/10

Psychomotor Vigilance Task

The PVT (Dinges & Powell, 1985) is widely used among scientists studying the effects of sleep restriction and circadian desynchrony because human performance on the task has proven to be highly sensitive to such factors (Dorrian et al., 2005). The task requires simple sustained attention to a high signal rate, typically for a period of 10 minutes. Thus, participants must maintain stable goal-directed alertness, which can be quite difficult when sleep deprived. During the task, participants are seated and visually fixed on a computer screen while holding a response box. Each time a red light stimulus appears in the window, participants respond by pressing the button on the box. The critical measure is how long it takes them to press the button after the stimulus appears. Stimuli appear randomly at 2-10 s ISI. Once the response button is pressed, the reaction time is displayed, to motivate the participant to respond as fast as possible without making errors of commission.

Results

The apparent simplicity of the PVT as a performance task belies the richness of the data and the complexity of the brain's response to sleep deprivation. Extensive research in the Dinges laboratory, where the PVT was developed, has revealed that responses to the PVT contain information on a number of facets of behavioral capability, and provide insights into the nature of neurobehavioral functioning in the presence of elevated biological pressure for sleep (Dorrian et al., 2005). Several phenomena of interest in PVT results are illustrated in Figure 2. The data are from an 88-hour, total sleep deprivation (TSD) study described in detail in Van Dongen (2004). Data on the PVT were collected in

10-minute test bouts performed every 2 hours. Subjects were awake throughout all days. The data in Figure 2 show baseline PVT performance and performance after 1, 2, and 3 days of TSD.
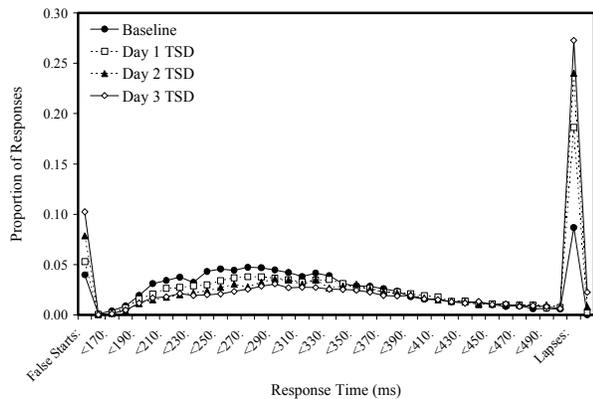

Figure 2: PVT data distributions for increasingly sleep deprived healthy adults.

The data in Figure 2 are presented according to how quickly or slowly participants responded. If participants responded during the delay period or within 150ms of the stimulus presentation, the response was characterized as a false start. As Figure 2 shows, the likelihood of a false start increases as sleep deprivation increases (Doran et al., 2001). Response times between 150ms and 500ms were considered to be in the normal alert range, and are plotted as a proportion of responses in each 10ms interval in Figure 2. As sleep deprivation increased, there was a shift toward longer response times in this range, with the fastest response times becoming less likely.

The second-to-last point on each of the lines in Figure 2 represents lapses. Lapses were defined a priori (by convention) as RTs greater than 500ms (but less than 30,000ms). As Figure 2 displays, there was a substantial increase in the proportion of responses classified as lapses (from approximately 8% at baseline to approximately 28% with 3 days of TSD). Finally, the last point on each line represents sleep attacks, where the participant failed to respond within 30,000ms of stimulus onset. This triggered an alarm to wake the participant. The increase in sleep attacks (from almost none to just over 2%) was smaller than the increase in lapses, but they represent a dramatic failure to respond.

These findings present significant modeling challenges. One unusual aspect of the data is the impact of sleep deprivation on false alarms (errors of commission), which increase at the same time that appropriate responses become slower and less likely (errors of omission). In addition, while response times in the normal range increased subtly, the large increase in lapses and the smaller increase in sleep attacks suggest that something more substantial is occurring than just a simple slowdown in processing – variability in performance increases as sleep deprivation increases (Doran et al., 2001). The ACT-R model described next illustrates the account of these effects that we have developed.

Cognitive Model

ACT-R is a production system, with a distinction between declarative memory, or memory for facts and information, and procedural memory, or memory for operations and transformations. In addition to these memory components, ACT-R has perceptual and motor modules that allow the system to interact directly with experimental software under realistic perceptual-motor constraints. Each module has an associated set of buffers (e.g., the visual system has a visual-object buffer and a visual-location buffer), which holds current information relevant to that module. ACT-R runs in cycles where the current state (i.e., the contents of the buffers) is compared to the conditions of the set of actions, and one is selected from the set of productions that match the current state. That production is then executed, or fired, and a new state emerges based on the consequences of the production. The selection of which matching production to fire is governed by an equation that calculates the "expected utility" (E) of using that production to reach the goal using the equation:

$$E = PG - C + \varepsilon$$

Here, P is the estimated probability of reaching the goal if the production is fired, and C is the expected cost (in seconds) of achieving the goal if the production is fired. G is an architectural parameter that has been loosely referred to as the value of the goal. The equation also incorporates a stochastic component, ε, which makes the calculation of E noisy. The production whose condition matches the current state of the system and has the highest value of E is selected and fired on each cognitive cycle[1]. This process is modulated by a utility threshold. If there is no production with an expected utility above the threshold, then no actions are performed. While this circumstance is usually avoided in ACT-R models, it is an important feature of this model, the details of which are described next.

Model Design

Due to the simplicity of the PVT, the model which performs it is relatively straightforward. The model is driven by procedural knowledge since the task simply requires a motor response to a visual stimulus. Using ACT-R's perceptual and motor modules, the model interacts directly with an implementation of the PVT. The model performs the task by waiting for the stimulus to appear, and responding with a button press.

There are two productions that may fire during the delay before the stimulus appears; a *wait* production and a *just-click* production. The *wait* production represents appropriate behavior during the inter-stimulus interval. In contrast, the

---

[1] The probability of selecting a specific production on a given cycle is governed by the Production Choice Equation in Anderson et al. (2004, p. 1044).

just-click production represents the capacity for a false alarm. This production executes a mouse-click regardless of whether or not there is anything in the model's visual buffers (i.e., regardless of whether or not ACT-R "sees" anything in the window on the monitor).

Once the stimulus appears, a sequence of two productions must fire to produce a response. The first shifts visual attention to the presented stimulus (i.e., recognizes that the stimulus has been presented) and the second executes a response. The *wait* production does not fire once the stimulus appears because it is deliberate inaction that depends on nothing being on the screen. However, because the *just-click* production does not consider whether or not something is on the screen, there is a chance it may fire at any point, whether or not the stimulus has been presented.

The brief description just provided leaves out one important component of the model, the utility threshold. It is possible for none of the productions to rise above threshold on a given cycle. If this happens, the system is idle for the length of that cycle (about 50ms). After the cycle, the expected utility values are recomputed, and if a production rises above threshold it is executed. The addition of noise to the calculation of *E* means that it is possible for nothing to happen on one cycle, followed by a cycle in which a production is executed.

The "empty" cognitive cycles represent the model slipping off to sleep while performing a task. However, this should mean that arousal is decreasing over time. To represent this in the model, a mechanism was added that decrements *G* when none of the matching productions rise above threshold after the stimulus has been presented. Each time this happens, the value of *G* is decreased by .035, which effectively lowers the utility values of the appropriate productions. The impact of this mechanism is that as time passes after the stimulus appears, it becomes less likely that the model will produce a response on any given cycle. In the current model, the value of *G* is reset at the beginning of each delay period. However, alternatives for calculating this value in an ongoing manner are being pursued. The initial values of *G* used here are discussed in the next section.

It is the interplay of the productions and the utility threshold that produces the model's behavior. During the inter-stimulus interval, the model can behave appropriately by waiting or doing nothing on each cycle. Doing nothing occurs when both the *wait* production and the *just-click* production have *E* values below threshold. This still results in appropriate behavior during the delay, because nothing should be done during that interval. Once the stimulus appears, the model may either *attend* (if the stimulus has not been attended), *respond* (if the stimulus has been attended), or do nothing. In this situation, doing nothing is no longer appropriate behavior, and it only occurs if none of the *E* values for eligible productions are above threshold. If it persists for long enough, the model produces a lapse. If doing nothing continues for a really long period, the model produces a sleep attack.

Parameters A total of three parameters in this model varied from their default values. Of these, two were not manipulated across different levels of sleep deprivation. The first of these is the utility threshold. The calculated value of *E* must be greater than this value for the production to fire on a given cycle. In the model presented here, the utility threshold was set to 1.75. The default setting is 0. However, most models in ACT-R maintain all productions above threshold and rely on the production matching process and conflict resolution to determine which production will fire. Thus, little research has been done examining the role of this parameter on performance in ACT-R.

The second parameter that was used in this model was the probability of success for the *just-click* production. This was set to 0, which results in a comparatively low *E*. This reflects the idea that random clicking is not likely to lead to success on the task. The net impact is to reduce the likelihood that the *just-click* production will fire relative to the *wait* production during the delay and relative to the *attend* and *respond* productions after the stimulus has been presented.

The third parameter was adjusted to produce the effects of sleep deprivation seen in the data. This parameter was *G*. While *G* has been referred to as the "value of the goal" in general, some researchers have viewed this parameter as reflecting "arousal" or "motivation" in studies of how stress or mental fatigue impact performance (Belavkin, 2001; Jongman, 1998). It is in this sense that *G* is conceptualized in this model.

By viewing *G* as a measure of arousal, it is natural to see one of the impacts of sleep deprivation as a lowering of *G*. This has the effect of decreasing the probability that expected utility will rise above threshold for any of the productions. In addition, the value of *G* plays a large role in determining the likelihood that a production will be executed, a role attributed to the thalamus in Figure 1. Because activity in the thalamus is impacted by sleep deprivation, *G* appears to be an excellent candidate for a parameter that may be impacted by lack of sleep.

The value of *G* was set to 1.87 for the baseline condition. As sleep deprivation increased, *G* was decreased, representing a decreased level of arousal. In this model, the value of *G* was set to best fit the observed data (though alternatives mechanisms are being explored). *G* was set to 1.77, 1.72, and 1.68 to represent the effects of 1, 2, and 3 days of TSD respectively. These initial values are decremented on empty cognitive cycles to represent the model drifting off to sleep.

Finally, there is a flag in ACT-R that allows cycle times and motor actions to be noisy. By default, these processes take a set amount of time (e.g., the default cycle time is 50ms). In our model, these times varied between 50% and 150% of their default values, according to a uniform distribution. This does not impact the behavior of the model. However, this mechanism is necessary to allow the model to

produce a continuous distribution of response times, rather than response times distributed at specific intervals.

Results

The changes in the *G* parameter have a large impact on the model's performance. The data from the model are presented in Figure 3. It illustrates that all but one (false alarms) of the major trends in the human data are closely matched by the model. The correlation to the human data is 0.986 and the RMSD is 0.55%. The data and predictions presented in Figures 2 and 3 are on the same scale, and the RMSD indicates that the overall correspondence between the two datasets is quite close.
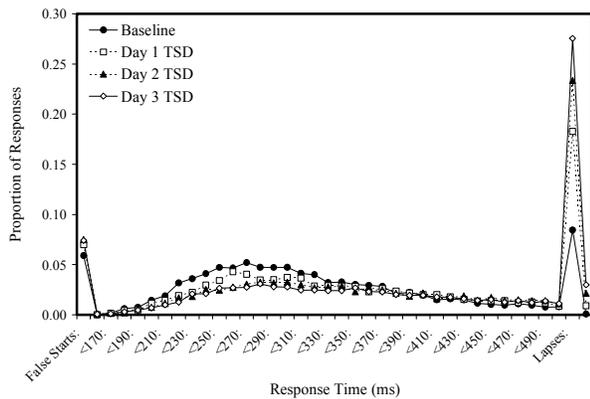

Figure 3: Model data from PVT.

As noted above, the model's performance results from the interaction of the utility values for the available productions and the utility threshold. When the stimulus is on the screen, either the *attend* or the *respond* production will match. However, on some occasions the production will fall below threshold, resulting in a temporary delay in the response. The likelihood of this happening increases as *G* is lowered because the expected utility value is lower before noise is added. In addition, because *G* is lowered when ACT-R does nothing on a cognitive cycle, it becomes increasingly unlikely that the model will respond at all each time the model misses an opportunity to respond. In effect, the model drifts off to sleep.

The manipulations of *G* produce the increase in lapses and sleep attacks, as well as the shift in normal response times in the model. Interestingly, the model does not require an increase in cycle time as a function of sleep deprivation. The shift in normal response times is due to the decreased probability of the model responding at the first opportunity as *G* is lowered. The immediate result is fewer fast responses, but the effects accumulate and translate into many more lapses and more sleep attacks, as occurs in the experimental data (Doran et al., 2001).

It is notable that during sleep deprivation, the model's false alarm rate increases. However, model false alarms did not show the degree of increase that was found experimentally, although the trend is in the right direction. The reason that false alarms increase in the model is that the

relative probability that the *just-click* production will fire actually increases as *G* decreases. Whereas *E* decreases for the *wait* production as *G* is lowered, *E* for the *just-click* production remains unchanged, because the probability of success for the *just-click* production was fixed at 0. Thus, with lower values of *G*, only the *wait* production has a reduced likelihood of firing, thereby making it a little more likely that *just-click* will fire. The result in the model is a subtle increase in false alarms as sleep deprivation increases. Further work is needed to understand the source of the comparatively large increase in false alarms shown by human participants.

Conclusions and Future Directions

The model presented in this paper provides the first successful approach to cognitive modeling of how sleep deprivation may lead to poorer performance on a task that requires sustained attention and rapid responses to frequent signals. PVT performance changes in a complex manner as sleep deprivation increases (Doran et al., 2001; Dorrian et al., 2005), which presents challenges to modeling these effects. The ACT-R approach used here appears to have captured most of the effects of sleep loss on PVT performance. As such, the model has increased our understanding of how the effects of fatigue from sleep deprivation can be captured in a formal cognitive model.

Our approach was to use existing research to narrow the range of possible mechanisms for producing fatigue effects in ACT-R. The neurophysiological work that has been done in the fatigue community and the recent work in ACT-R on mapping architectural components to brain areas, appear to provide useful constraints for identifying which parameters to manipulate in order to capture the effects of sleep loss. By combining these research areas, we were able to identify a parameter in ACT-R that (1) influences the behavior of our model, with outcomes that are very similar to those found in human performance before and during sleep derivation, and (2) corresponds to a neural structure (i.e., thalamus) that has been identified as influential in the brain's response to sleep restriction (Portas et al., 1998).

The thalamus is linked to the production execution component of ACT-R. Manipulations in *G* influence the likelihood that a production will be executed, by influencing the probability that expected utility (*E*) will rise above threshold. This suggests that changes in activity in the thalamus of human participants may be related to changes in *G* in the ACT-R architecture. The model described here demonstrates that such an account is feasible by using *G* to produce performance differences in the model that closely match changes in human performance as a result of sleep deprivation.

We are currently extending this research to other simple tasks that are used in sleep restriction research. This work seeks to identify a small set of ACT-R parameters, corresponding to a small set of neural structures, that are impacted by sleep deprivation. Once these parameters are identified, we will use cognitive performance data and the

biomathematical models that capture the temporal dynamics in these data to drive changes in their values. The resulting theory will then be used to make *a priori* predictions about human performance on complex dynamic tasks under conditions of degraded cognitive functioning.

Acknowledgments

References

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y . (2004). An integrated theory of the mind. *Psychological Review 111*, 1036-1060.

Belavkin, R. V. (2001). The role of emotion in problem solving. In C. Johnson (Ed.), *Proceedings of the AISB '01 Symposium on emotion, cognition, and affective computing* (pp. 49-57). York, England: Heslington.

Caldwell, J. A., Caldwell, J. L., Brown, D. L., & Smith, J. K. (2004). The effects of 37 hours of continuous wakefulness on the physiological arousal, cognitive performance, self-reported mood, and simulator flight performance of F-117A pilots. *Military Psychology, 16*(3), 163-181.

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., & 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., Dinges, D. F (2001). Sustained attention performance during sleep deprivation: evidence of state instability. *Archives of Italian Biology: Neuroscience 139,* 3, 253-267.

Dorrian, J., Rogers, N. L., & Dinges, D. F. (2005). Psychomotor vigilance performance: Neurocognitive assay sensitive to sleep loss. In Kushida, C.A. (Ed.), *Sleep Deprivation: Clinical Issues, Pharmacology and Sleep Loss Effects*. Marcel Dekker, Inc., New York, NY, pp 39-70.

Durmer, J. S., & Dinges, D. F. (2005). Neurocognitive consequences of sleep deprivation. *Seminars in Neurology* in press.

Drummond, S. P. A., & Brown, G. G. (2001). The effects of total sleep deprivation on cerebral responses to cognitive performance. *Neuropsychopharmacology, 25*, S68-S73.

Jongman, L. (1998) How to fatigue ACT-R? In *Proceedings of the Second European Conference on Cognitive Modelling* (pp. 52-57). Nottingham: Nottingham University Press.

Lin, J. S. (2000). Brain structures and mechanisms involved in the control of cortical activation and wakefulness, with emphasis on the posterior hypothalamus and histaminergic neurons. *Sleep Medicine Reviews, 4*, 471-503.

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 & Environmental Medicine 75,* 3, A4-A14.

Moruzzi G, Magoun HW. (1949). Brainstem reticular formation and activation of the EEG. Electroencephalography and Clinical Neurophysiology, 1, 455–473.

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. *The Journal of Neuroscience, 18*, 8979-8989.

Ritter, F. E., Reifers, A., Klein, L. C., Quigley, K., & Schoelles, M. (2004). Using cognitive modeling to study behavior moderators: Pre-task appraisal and anxiety. In *Proceedings of the Human Factors and Ergonomics Society* (pp. 2121-2125). Santa Monica, CA: Human Factors and Ergonomics Society.

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), 15-36.