Computational Modeling of the Combined Effects of Circadian Rhythm and Sleep Deprivation

Joshua B. Gross (jgross@ist.psu.edu)
College of Information Sciences and Technology, The Pennsylvania State University
University Park, PA 16801 USA

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 (hvd@wsu.edu)
Sleep and Performance Research Center, Washington State University Spokane
P.O. Box 1495; Spokane, WA 99210-1495, U.S.A.

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 USA

Abstract

Previous modeling efforts have demonstrated a set of ACT-R mechanisms that match the effects of fatigue caused by sleep deprivation in humans. This paper describes an extension of that research, which involved integrating biomathematical models of alertness with this architectural representation of how fatigue affects cognition. We used the newly integrated models to account for the combined effects of circadian rhythm and sleep deprivation on cognition as it relates to performance in a sustained-attention task. Using this system, we were able to accurately reproduce changes over time in a set of key performance measures. This represents a significant advance toward the goal of developing models that predict performance when the cognitive system is fatigued.

Introduction

Performance in task environments is typically assumed to be at normative levels; this assumption is made both in and out of the laboratory. However, there is an increasing awareness within the cognitive science community that many (often subtle) factors can affect performance, including emotion (Gratch & Marsella, 2004; Hudlicka, 2003), motivation (Belavkin, 2001), stimulants (Ritter, *et al.,* 2004), and fatigue (Gunzelmann, *et al.*, 2005). These factors, generally called *cognitive moderators*, provide a deeper understanding of the sources of variance in human performance. They also give rise to a more complete view of the multi-faceted nature of cognition and a more holistic perspective on the total cognitive system. This paper reports progress toward predictive models of the effects of fatigue[1] on cognition.

Fatigue and Human Error

There are many documented cases in which restricted sleep or total sleep deprivation have been implicated as a cause of human error that led to accidents. Famous examples include the grounding of the Exxon Valdez and the reactor meltdown at Chernobyl (Caldwell, 2003). These examples, while dramatic, highlight a key concern in this paper: sleep deprivation can have drastic effects on performance.

Despite the potential for negative consequences, occasional sleep deprivation and insufficient sleep for long periods are widespread problems in our society. For instance, despite regulations requiring minimum rest periods, truck drivers often fail to obtain adequate sleep (Dinges, 1995). Military missions frequently require extended wakefulness due to distance or complexity, and nighttime operations have become the norm rather than the exception (Caldwell, 2003). Computational models that make *a priori* predictions regarding the performance consequences of sleep loss can be used to help mitigate human error from fatigue, and can thus contribute to improved safety and optimal readiness to perform.

Predicting Performance Under Sleep Deprivation

Predicting performance under sleep deprivation is important for anyone attempting to design schedules that maximize human usefulness under resource limitations. Furthermore, researchers in human-computer interaction (HCI) and human factors (HF) can use performance predictions when designing equipment or software likely to be used under fatigued conditions. Extensive laboratory research into fatigue has produced large amounts of human performance data. These data provide a good foundation for developing and evaluating methods for modeling fatigued performance, as illustrated in the present paper.

Biomathematical Models A number of biomathematical models of fatigue have been developed within the biomedical research community (see Van Dongen, 2004). These models are available in software tools that allow the user to construct a protocol consisting of periods of sleep and wakefulness. Using this protocol, one can generate

[1] In this paper, lack of sleep is equated with the induction of fatigue; space limitations prohibit discussion of more subtle aspects of this relationship.

estimates of relative cognitive functioning, or alertness. We chose to use two such models, in order to compare their relative efficacy and accuracy, and to provide flexibility for future research. These models are the Circadian Neurobehavioral Performance and Alertness (CNPA) (Jewett & Kronauer, 1999), and the Sleep, Activity, Fatigue, and Task Effectiveness (SAFTE) (Hursh *et al.*, 2004).

The alertness measures produced by CNPA and SAFTE can be seen as an inverse measure of fatigue. The models combine sleep-loss-induced decreases in performance with a cyclical model of circadian rhythm (see Figure 1). Van Dongen (2004) reviewed a number of such models and showed that they capture some of the important dynamics associated with fatigue under conditions of total sleep deprivation. However, the models cannot make predictions about performance on specific tasks. The alertness measures produced can be scaled or transformed to fit particular datasets, but this is necessarily a *post hoc* process.

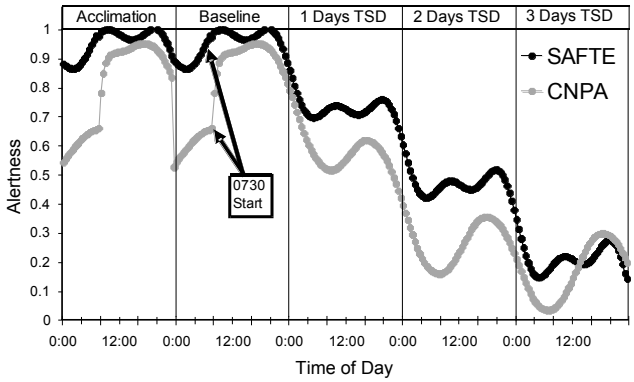

Figure 1: Biomathematical model predictions of alertness, across 88 hours of total sleep deprivation (TSD). The 88-hr period began at 7:30 AM on the baseline day, as indicated in the Figure.

Computational Models The long-term goal of this research project is to provide a comprehensive computational account of the effects of fatigue on performance; this will allow *a priori* predictions of the impact of sleep loss in specific tasks. Performance changes can be seen in process measures (such as response times (RT) and errors), as well as overall task outcome measures (i.e. successful completion). Prior research (Gunzelmann *et al.*, 2005) established preliminary aspects of this method, identifying a neurobehaviorally-inspired approach to fatiguing the ACT-R cognitive architecture. This approach was validated against human data at a temporal resolution of whole days of sleep deprivation (` measures averaged within each day).

This paper describes an extension of the earlier research, which adopts a fusion of modeling approaches, capitalizing on the strengths of both biomathematical fatigue models and computational cognitive models. We used the alertness predictions from the biomathematical models to drive changes to architectural parameters in ACT-R (described below). This process produced model performance changes that very closely replicated human performance changes, illustrating the ability to account for changes in performance that occur over 88 hours of total sleep deprivation (TSD) due to the combined impact of sleep loss and circadian rhythm. In addition to fitting the data, we illustrate how biomathematical models, like those that have been developed to describe the effects of fatigue, can be integrated into a cognitive architecture, to expand the explanatory power of both systems.

Human Subject Test Protocol

Van Dongen and Dinges (2005) reported the empirical study that produced the human data used here. Participants were brought into the lab for three days of acclimation and baseline recordings, where they were given 8 hours time in bed per night (23:30 to 07:30). Beginning at 07:30 after the third night in the lab, participants were kept awake continuously for 88 hours. During all waking periods, participants performed a 30-minute battery of computer-based tests and questionnaires every two hours.

The focus of this paper is on data from one of the computer tests, the psychomotor vigilance task (PVT; Dinges & Powell, 1985). The PVT is a sustained-attention task that requires the participant to monitor a computer screen for a stimulus, which appears at a known location but at a random interval between 2 and 12 seconds. The participant's task is to press a response button as soon as the stimulus appears. The critical measure is the latency of that response, i.e., the reaction time. Participants are instructed to respond as quickly as possible, while avoiding false starts. In addition to sustained attention, this task captures performance for a general class of tasks that require both vision and manual action, and for which reaction time (RT) is crucial. In the Van Dongen and Dinges (2005) study, 13 participants performed a 10-minute session of this task every 2 hours. Across the 88 hours of sleep deprivation, then, all participants completed 44 of these sessions, providing a rich source of data on this task.

Following conventions in the fatigue research community, reaction time data are characterized as follows: false starts (button presses before the stimulus appears or within 150 ms of stimulus onset), alert responses (RTs from 150 ms to 500 ms after stimulus onset), lapses (RTs greater than 500 ms but less than 30 s), and sleep attacks (no response within 30 s of stimulus onset).

Computational Cognitive Model

The computational model described here is implemented in the ACT-R cognitive architecture (Anderson *et al.*, 2004). The critical components of the ACT-R model for the PVT include the perceptual and motor modules, as well as the central production system. The perceptual and motor modules in the architecture allow ACT-R to interact directly with software implementations of experimental tasks, incorporating realistic timing constraints on those operations. In a task like the PVT this is vital, since performance is almost entirely dependent on perceptual-motor processes.

Cognition in ACT-R is represented by the serial execution of productions. At any point, the state of the system is

represented by the contents of a set of buffers, which serve as the interface between peripheral modules (like the perceptual and motor modules), and the central production system. Productions match against the contents of those buffers. When a matching production is executed (fired), it serves to modify the contents of the buffers directly, or to make requests of particular modules (e.g., to act or observe), which result in changes to buffer contents (thus producing a new state).

Mechanisms for Fatigue

The central production system is the component of the architecture that was targeted for defining mechanisms for fatigue effects based on previous research in ACT-R and on neurobehavioral research on the effects of fatigue. In this section we describe those mechanisms. Gunzelmann *et al.* (in press) provide a more detailed discussion of the empirical and theoretical motivations underlying this implementation of fatigue in ACT-R.

In the production system in ACT-R, productions are matched against conditions (i.e., buffer contents) and one is selected and fired, which generally produces some change. Production selection is controlled by the following equation, which is used to calculate a utility value (*U)* for each production:

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

In this equation, P is the probability of achieving the goal if that production is used (by default, P=1), and C is the anticipated cost. ε is a noise parameter that produces stochasticity in the selection process. G is a parameter that has been cast as "motivation" or "arousal" (Belavkin, 2001; Jongman, 1998), and we conceptualize it as a representation of alertness.

In our approach, the value of G is decreased to represent lower levels of alertness within the architecture. This has the effect of lowering the utility value for any production where P>0. As described below, the initial value of G for each session is estimated using the biomathematical models of fatigue.

A utility value (*U)* is calculated for each production that matches the current state (that is, every viable production). The production with the highest value for U is selected. This production is executed, provided that the value for U exceeds the utility threshold, T_u. The value for T_u also varies in the model, representing attempts to compensate for fatigue. As alertness decreases, the value of T_u decreases, which makes it easier for productions to exceed the threshold and fire. In cases where no production exceeds T_u, no cognitive actions occur on that cycle, producing a "micro-lapse" lasting for approximately 50 ms. Decreasing G makes micro-lapses more likely to occur. An increasingly long series of micro-lapses produces longer reaction times in the PVT (including RTs categorized as lapses, as well as sleep attacks).

The last mechanism in the model for representing the impact of fatigue relates to the G parameter. As noted, we take this value to represent alertness. The micro-lapses that occur when no productions exceed T_u are indicative of decreasing alertness (falling asleep). As alertness decreases, the likelihood of an inactive cycle increases. To capture this phenomenon, the value of G is decremented on cognitive cycles where no productions exceed T_u. Each time this occurs, the value of G is reduced by .035[2]. As a result, the model becomes progressively less likely to execute an action. The noise value used in the utility calculation introduces stochasticity in each cycle. The value of G is restored to its initial value (estimated from the biomathematical models) at the beginning of each simulated trial.

The most important aspect of this approach is that it makes full use of ACT-R's subsymbolic computational layer to generate fatigue. This has two significant advantages. First, the mechanisms can be generalized to other ACT-R models, for other tasks (although parameter values and additional controls may be necessary). Second, the knowledge within the model itself is not modified to simulate fatigue; this simplifies model development. It also reflects the more cognitively plausible explanation that the effects of fatigue occur at the architectural level of cognition, rather than the knowledge or symbolic level. While fatigue may result in changes in how knowledge is used (e.g., strategy shifts), it does not cause changes in the knowledge itself.

Model Dynamics

The fatigue mechanisms just described interact with default ACT-R mechanisms and the knowledge incorporated into the model to produce the task performance discussed below. At any point in the PVT task, the model has three options available (italics indicate productions in the model). First, the model may behave appropriately, explicitly *waiting* during the delay interval, then *attending-to* and *responding-to* the stimulus once it appears.

Another production represents the capacity to *just-respond*, regardless of whether the stimulus has been presented or not. This production can fire at any point in the task, producing false starts[3] in cases when it fires before the stimulus appears or within 150 ms of stimulus onset. Because it is unlikely to result in a correct response, it is given a probability of success (P) of 0[4]; thus the utility of the *just-respond* production is immune to changes in G.

Finally, the model may fail to execute any cognitive action at all on any given cycle. If neither the appropriate action nor the *just-respond* production have values of U that exceed T_u, then no productions are fired on that cycle, and G is decremented as described above. With lower values of G, this becomes more likely. The result is longer reaction times and increasing proportions of responses that are classified as lapses or sleep attacks.

[2] This G-reduction parameter is a new architectural claim. There is no previously existing default or common value for this parameter. Its value was selected to best fit the observed data.

[3] An inhibition-lowering approach to generate false starts may be more justifiable, but architectural limitations make this infeasible at present, and it would not significantly alter the results.

[4] While the probability of success is not in fact 0, a number slightly above 0 would produce the same results.

An appropriate side effect of decreasing G is that the architecture becomes increasingly likely to execute low-cost alternatives, focusing less on the probability of success. In the model, this produces a higher probability of committing a false start. This is because decreased values for G serve to diminish differences in U between the appropriate action (where P=1) and the inappropriate action (P=0).

Using Alertness Predictions to Drive Parameters

The goal of this research is to enable predictive modeling of the effects of fatigue. Since the biomathematical models make *a priori* characterizations of global changes in performance as a function of sleep loss, it makes sense to utilize those predictions to control parameter values for the fatigue mechanisms that have been introduced in ACT-R.

Our research hypothesis was that, if the fatigue levels accurately measure relative performance, then a simple linear scaling should be possible to map alertness to the fatigue-controlling parameters of ACT-R (described above). As noted, decreasing G results in a lower likelihood of any production firing; this is consistent with increasing fatigue. Also, because individuals are motivated to maintain performance levels in these protocols (and are provided with trial-by-trial performance feedback in the PVT), we include a decrease in the utility threshold (T_u) to represent compensatory behavior. This allows lower utility values to result in the execution of a production in the model.

To establish an initial link between parameter values in ACT-R and alertness measures in the biomathematical models, we identified the session during the experiment where alertness measures were highest and the session where they were lowest. For each of these, we identified the best fitting values for G and T_u, We used those values to set the high and low boundaries for G and T_u. Finally, we combined the alertness for each point into a linear scaling equation that produced an ACT-R parameter value:

$$G_t = A_t \cdot G_r + G_{\min}$$

and

$$Tu_t = A_t \cdot Tu_r + Tu_{\min}$$

where G_t is the value for G at time t, A_t is the calculated alertness for time t (from either CNPA or SAFTE), G_r is the range of G ($G_{max} - G_{min}$), and G_{min} and G_{max} are the minimum & maximum values for G. The second formula is identical, except substituting the T_u for G in inputs and outputs. Again, our minima and maxima for these parameters come from identifying the best-fitting parameters for the sessions where the biomathematical predictions of alertness were highest and lowest. The values used in the fits described below are presented in Table 1.

Table 1: Parameter values used to calculate G and T_u.

Parameter	G	T_u
Minimum	1.54	1.68
Maximum	2.02	1.88
Range	.48	.2

Once the predicted values for G and T_u were calculated for each given time in the experiment, we ran the model using those parameter values. The performance predictions from these model runs (based on 100 repetitions of a 10-minute PVT session) were then compared to the averaged human participant data, as described in the next section.

Model Performance

As noted above, the human participants were sleep-deprived for 88 hours and tested every 2 hours during that time, resulting in 44 test points. We compared the human performance data for each session with predictions from the ACT-R model, using both CNPA and SAFTE estimates of alertness. The performance measures included false starts, lapse frequency, median alert reaciton times, and sleep attacks. Figures 2-5 compare the performance of the ACT-R model to human participant performance for each of these dependent measures for both models of alertness. Correlations and RMSDs for each of the dependent measures are presented in Table 2.

Table 2: Quantitative comparison of ACT-R model predictions to human data.

Dependent Measure	ACT-R using CNPA		ACT-R using SAFTE	
	r	RMSD	r	RMSD
False Starts	.71	.022	.81	.019
Median Alert RT	.70	17.82	.67	16.57
Lapses	.88	.055	.83	.060
Sleep Attacks	.72	.014	.83	.026

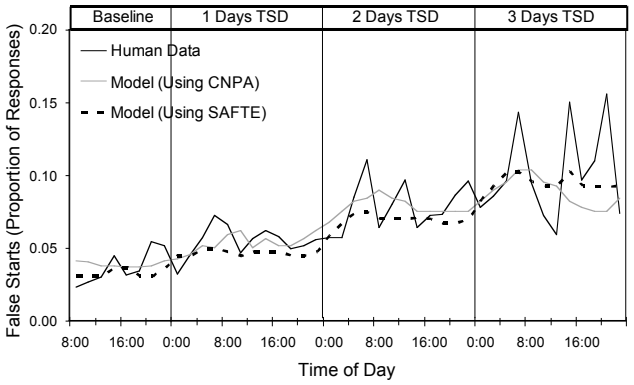

Figure 2: Human data and ACT-R predictions of false starts based on CNPA and SAFTE across 88 hours TSD.

Circadian rhythm actually becomes more pronounced as sleep deprivation continues. This is illustrated in Figure 1 by the alertness predictions and is borne out in the human performance data seen in Figures 2-5. For all performance measures, we see a precipitous fall (in both model and empirical data) at the expected nadir (early morning), followed by an improvement in performance. In fact, the worst performance (in most performance measures) is not, as one might expect, near the end of the experiment, but

early in the morning on the last day; this is approximately 16 hours before the maximum amount of sleep deprivation.

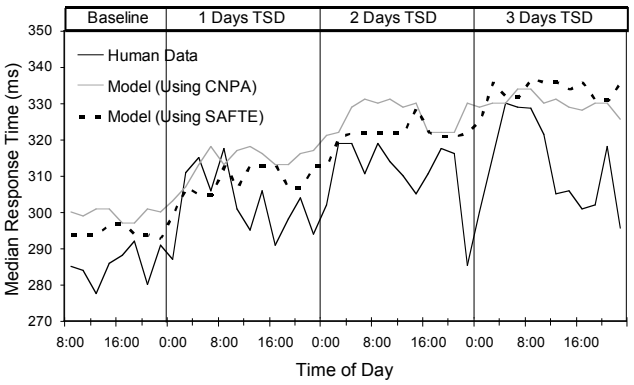

Figure 3: Human data and ACT-R predictions of median alert reaction times based on CNPA and SAFTE across 88 hours TSD.

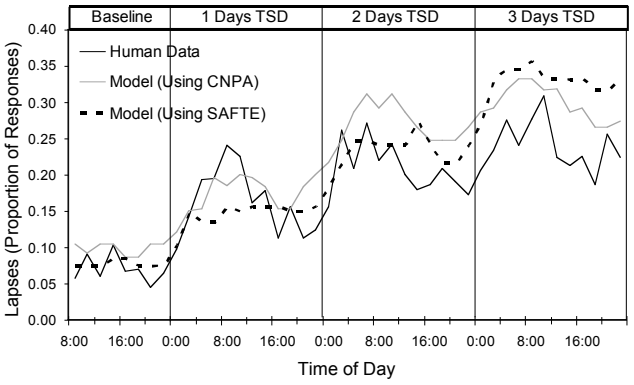

Figure 4: Human data and ACT-R predictions of lapses based on CNPA and SAFTE across 88 hours TSD.

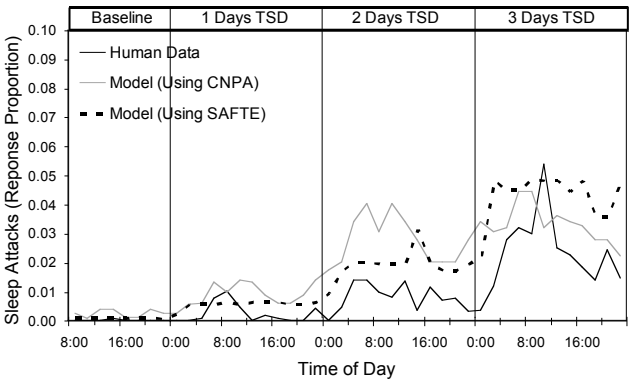

Figure 5: Human data and ACT-R predictions of sleep attacks based on CNPA and SAFTE across 88 hours TSD.

Figure 4 shows the ability of our approach to predict lapses accurately. This can been seen in the high correlation and low RMSD for lapses shown in Table 2. The trends for false starts (Figure 2) and median reaction time (Figure 3) are also captured by the model. These results illustrate the importance of having a model that actually performs the task. Biomathematical models must be fit to each of the dependent measures independently, while the model makes relatively accurate predictions across dependent measures using the same model with the same parameter values.

Sleep attacks (Figure 5) turned out to be more difficult to predict, due in part to a lack of sleep attacks early in the experiment, followed by a large increase late in the experiment. Although the model captured the overall increase and the qualitative trend fairly well, the paucity and irregularity of sleep attacks make this performance measure specifically susceptible to model misfit resulting merely from stochastic variability. This is exacerbated by the relatively small sample size (N=13).

Overall, these results indicate that it is possible to predict performance on the PVT with a relatively high degree of accuracy. The model successfully captured the performance of human participants on this task, with a degree of detail that has not been presented in any other attempt at modeling PVT data under conditions of fatigue.

Conclusion and Future Directions

The mechanisms in the model that generate the effects on performance may seem, at first glance, to be complicated. They involve a cognitive architecture (ACT-R), a task model (PVT), two ACT-R parameters (G and T_u), and a biomathematical model of fatigue (CNPA or SAFTE). However, this sophisticated approach is necessary in order capture the dynamics of performance with fatigue as a moderator. In addition, all of the components of this approach are supported by empirical and theoretical evidence (Gunzelmann *et al.*, in press).

There are limitations to the model described here and there remains work to be done to improve on it. For instance, although the model does replicate the general ebb and flow of performance, it does not always capture smaller-grained variations, which exposes current limitations in mathematical models of alertness. In addition, theoretically grounded values and ranges for parameters (including G-decrement) will need to be determined in the future. Despite these limitations, this model represents a significant advance toward predictive modeling of the effects of fatigue on cognition. The biomathematical models of fatigue predict values for the control parameters in ACT-R, while ACT-R provides us with an implementation of the mechanisms of human cognition, complete with subsymbolic mechanisms that allow us to control aspects of the processing that occurs, By combining these approaches we have created a more complete, mechanistic explanation of how sleepiness moderates cognitive effectiveness.

This research extends and provides additional validation for the model of fatigued ACT-R produced by Gunzelmann *et al.* (2005; in press). The model is able to accurately

reproduce the circadian ebb and flow of human performance, not merely composite performance over daylong periods. The mechanisms that have been developed are the initial step in achieving a larger research goal of predicting the effects of fatigue on human task performance. This work also extends and helps to validate the ACT-R theory and architecture, and the biomathematical models of fatigue as well.

A critical contribution of this research is the integration across several disciplines of research to support the overall account. ACT-R could not simply be combined with the biomathematical models to produce these predictions. A theory and method for integration was also necessary. The different disciplines of origin of the tools and theories used (ACT-R, biomathematical fatigue models, neurobehavioral research) might suggest a potential for incompatibility. However, integrated approaches are imperative for cognitive science precisely because of its interdisciplinary foundations. The successful integration of these tools and methods also gives credence to a congruence between the respective underlying theories. This work fits well into the ongoing mission of cognitive science to study various aspects of cognition from the perspective of multiple disciplines.

From an applied perspective, the success of this research so far is encouraging. The ability to reproduce much of the variance in human performance under conditions of total sleep deprivation suggests that it may be possible to make informed decisions about sleep protocols with fewer expensive human subject experiments. Further work will extend this approach to individual performance, additional tasks, and different sleep deprivation and restriction protocols (see Van Dongen *et al.*, 2003). One issue we wish to address with the current method is that it requires data to establish the relationship between the ACT-R parameter *G* and alertness as predicted by the biomathematical models. The only way to acquire bottom-end data is to perform controlled studies, and as part of the value for a predictive model comes from obviating such studies, further work on this problem will be necessary.

Acknowledgments

Cognitive model development was sponsored by the Air Force Research Laboratory's (AFRL) Warfighter Readiness Research Division and by grant number 04HE02COR from the Air Force Office of Scientific Research (AFOSR). Human data collection was sponsored partly by AFOSR grants F49620-95-1-0388 and F49620-00-1-0266, and by NIH grant RR00040. Joshua Gross's participation was sponsored by an AFRL internship administered by the Oak Ridge Institute for Science and Education.

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. (2003, Fall). Wake up to the importance of sleep for air safety! *Flightline*, 30-33.

Dinges, D. F. (1995). An overview of sleepiness and accidents. Journal *of Sleep Research, 4(2)*, 4-11.

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.

Gratch, J., & Marsella, S. (2004). A domain-independent framework for modeling emotion. *Cognitive Systems Research, 5*, 269-306.

Gunzelmann, G., Gluck, K. A., Price, S. C., Van Dongen, H. P. A., & Dinges, D. F. (in press). Incorporating the effects of sleep deprivation into a cognitive architecture. *Integrated Models of Cognitive Systems*. New York: Oxford University Press.

Gunzelmann, G., Gluck, K. A., Van Dongen, H. P. A., O'Connor, R. M., & Dinges, D. F. (2005). A neurobehaviorally inspired ACT-R model of sleep deprivation: Decreased performance in psychomotor vigilance. *Proceedings of the Twenty-Seventh Annual Meeting of the Cognitive Science Society* (pp. 857-862). Mahwah, NJ: Lawrence Erlbaum Associates.

Hudlicka, E. (2003). Modeling effects of behavior moderators on performance: Evaluation of the MAMID methodology and architecture. In *Proceedings of the 12th Conference on Behavior Representation in Modeling and Simulation* (pp. 207-215). Orlando: Institute for Simulation and Training.

Hursh, S. R., Redmond, D. P., Johnson, M. L., Thorne, D. R., Belenky, G., Balkin, T. J., *et al.* (2004). Fatigue models for applied research in warfighting. *Aviation, Space, and Environmental Medicine, 75*(3), A44-60.

Jewett, M. E., & Kronauer, R. E. (1999). Interactive mathematical models of subjective alertness and alertness in humans. *Journal of Biological Rhythms, 14*, 588-597.

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.

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*. Human Factors and Ergonomics Society, Santa Monica, CA.

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)*, A15-A36.

Van Dongen, H. P. A., & Dinges, D.F. (2005). Sleep, circadian rhythms, and psychomotor vigilance. *Clinical Sports Medicine, 24*, 237-249.

Van Dongen, H.P.A., Rogers, N.L., & Dinges, D.F. (2003). Sleep debt: Theoretical and empirical issues. *Sleep and Biological Rhythms, 1*, 5-13.