Using Computational Cognitive Modeling to Predict Dual-Task Performance With Sleep Deprivation

Glenn Gunzelmann, Air Force Research Laboratory, Mesa, Arizona, **Michael D. Byrne,** Rice University, Houston, Texas, **Kevin A. Gluck,** Air Force Research Laboratory, Mesa, Arizona, and **L. Richard Moore Jr.,** Lockheed Martin, Mesa, Arizona

Objective: The effects of fatigue on multiple-task performance were explored through computational cognitive modeling. **Background:** Fatigue typically has a negative impact on human performance. Biomathematical models exist that characterize the dynamics of human alertness, but the link between alertness and in situ performance on specific tasks is tenuous. Cognitive architectures offer a principled means of establishing that link. **Method:** We implemented mechanisms for fatigue, which produce microlapses in cognitive processing, into an existing model, adaptive control of thought–rational, and validated the performance predictions with Bratzke, Rolke, Ulrich, and Peters' data on fatigue and multiple-task performance. **Results:** The microlapse model replicated the human performance results very well with zero free parameters, although the fit was improved when we allowed two individual differences parameters to vary. **Conclusion:** Increased frequency of microlapses as a result of fatigue provides a parsimonious explanation for the impact of fatigue on dual-task performance and is consistent with previous research. **Application:** Our results illustrate how using biomathematical models of fatigue in conjunction with a cognitive architecture can result in accurate predictions of the effects of fatigue on dual-task performance. Extending and generalizing this capability has potential utility in any safety-critical domain in which fatigue may affect performance.

Cognitive impairments resulting from fatigue have critical implications in applied settings. Explanations for numerous industrial and commercial disasters, not to mention more common tragedies, such as car accidents (Gallup, 2002; Horne & Reyner, 1999; Klauer, Dingus, Neale, Sudweeks, & Ramsey, 2006; National Transportation Safety Board, 1995; Pack et al., 1995), have implicated fatigue as a contributing factor (Caldwell, 2003; Dinges, 1995; Mitler et al., 1988). Consequently, much of the research conducted on fatigue is directed at understanding the decrements arising from restricted sleep, extended periods of wakefulness, and circadian rhythmicity. An overarching goal is to come to an understanding of how sleep affects human cognition that is useful for making predictions about human performance in applied settings and thereby assist in planning and decision making (Dinges & Kribbs, 1991; Klerman & St. Hilaire, 2007). Because of the potential value of the research in these contexts, there is a long history of empirical and theoretical research, despite the cost and logistical issues associated with conducting controlled studies of fatigue.

In order to use research results from sleep studies to make predictions about task performance, one must understand two theoretical aspects of the issue. One necessary component is an accurate understanding of the endogenous and exogenous factors that influence overall level of alertness and how they interact to produce cognitive performance. To address this issue, biomathematical models have been developed that integrate mechanisms for time awake and circadian rhythms in estimating an overall level of cognitive functioning, or *alertness* (see Klerman & St. Hilaire, 2007; Mallis,

Address correspondence to Glenn Gunzelmann, Air Force Research Laboratory, 711 HPW/RHAC, 6030 S. Kent St., Mesa, AZ 85212-6061; glenn.gunzelmann@us.af.mil. ***HUMAN FACTORS,*** Vol. 51, No. 2, April 2009, pp. 251-260. DOI: 10.1177/0018720809334592.

Mejdal, Nguyen, & Dinges, 2004). These models embody the current understanding of the dynamics of alertness and some of the internal—physiological or biological—influences on overall cognitive performance.

Biomathematical models have been incorporated into tools that can be used to create and assess work–rest schedules (Dean, Fletcher, Hursh, & Klerman, 2007; Hursh, Raslear, Kaye, & Fanzone, 2006). Despite their sophistication, however, the models are limited to making general predictions about relative alertness. Their output can be scaled to existing human performance data to fit particular measures, but this must be done in a post hoc manner (e.g., Van Dongen, 2004) because the models predict relative alertness, not cognitive performance and behavior. This undermines a key application goal for such tools, which is to make predictions about the consequences of fatigue in particular task contexts when behavior and performance data are unavailable in real-world settings.

The limitations of biomathematical models point to the other necessary component of a predictive model of fatigue. Specifically, in addition to understanding the dynamics of performance overall, it is also necessary to understand the mechanisms through which fatigue operates to affect cognitive processing and, ultimately, behavior and performance. Among the theoretical accounts that have been developed in the sleep research community, a common perspective is that fatigue has two primary impacts on cognitive performance: a generalized *cognitive slowing* that affects performance in short-reaction-time tasks combined with an increased probability of *cognitive lapses*, which refers to breakdowns in cognitive processing lasting on the order of seconds (Dinges & Kribbs, 1991; Doran, Van Dongen, & Dinges, 2001; Heuer, Kleinsorge, Klein, & Kohlisch, 2004). However, these accounts are merely descriptive; they have little to say about the cognitive mechanisms underlying fatigue and how those mechanisms interact with others to produce decrements in cognitive performance.

MODELING APPROACH

Our research is positioned at the intersection of relative alertness predictions and quantitative performance predictions. Specifically, we use computational cognitive models that generate quantitative performance predictions in specific task contexts. The models are implemented in a cognitive architecture, which represents a theory of the information-processing mechanisms that underlie human cognitive capabilities (Newell, 1990). By using a cognitive architecture, one can implement theories of fatigue as influences on those computational mechanisms that lead to identifiable changes in the behavior of models. This involves manipulating parameters in the cognitive architecture related to aspects of its information-processing activity. Subsequently, when one links the dynamics of those parameters to the biomathematical models of alertness described earlier, it is possible to create a bridge between understanding the dynamics of fatigue and understanding the impact of those dynamics on cognitive processing and human behavior.

We have already demonstrated the value of this integration for generating principled, quantitative accounts of fatigued performance (e.g., Gunzelmann, Gluck, Kershner, Van Dongen, & Dinges, 2007; Gunzelmann, Gross, Gluck, & Dinges, in press). In this article, we explore extending such accounts to novel task contexts with the goal of making principled, a priori predictions.

The cognitive architecture that we are using is ACT-R, or *adaptive control of thought–rational* (Anderson et al., 2004). ACT-R is a computational theory of human cognition and performance that has been used to simulate human behavior in a wide variety of tasks (see Anderson et al., 2004, and Anderson & Lebiere, 1998, for partial reviews). ACT-R consists of multiple modules (e.g., cognitive, motor, visual) that can act in parallel. These modules communicate with central cognition through a set of buffers containing information about each module's processing (e.g., the visual module has two buffers that hold information about what is currently seen and where it is).

Central cognition in ACT-R is represented as a production system in which the patterns instantiated across the set of buffers lead to cognitive, perceptual, or motor actions. Many of ACT-R's modules have serial constraints; most relevantly, the procedural module in ACT-R can execute only a single production on a given

cognitive cycle, which represents a central bottleneck in cognitive processing. Additionally, all of the symbolic architectural mechanisms have associated numerical quantities that allow for continuous, graded changes in cognitive performance, a key feature of the degradations associated with fatigue. Further details about ACT-R are available elsewhere (e.g., Anderson, 2007).

The biomathematical model we use in the research described here is referred to as the *circadian neurobehavioral performance and alertness model,* or CNPA (Jewett & Kronauer, 1999). It produces an estimate of overall cognitive performance, identified as *cognitive throughput* in the model. The predictions take into account the influences of a sleep homeostat system and circadian rhythms, instantiating the widely accepted two-process model of human alertness (Achermann, 2004; Borbely, 1982). In addition, CNPA has mechanisms that represent the impact of light on the phase and amplitude of the circadian pacemaker and a sleep inertia process, which accounts for reduced levels of alertness when individuals initially awaken. The model has been shown to capture the dynamics of alertness as reflected by a variety of behavioral and physiological measures (e.g., Jewett & Kronauer, 1999).

We have used previous research to establish a link between biomathematical model predictions of alertness and parameters in ACT-R (e.g., Gunzelmann et al., 2007; in press). The results have shown that a simple linear mapping of alertness predictions to specific parameter values in ACT-R produces changes in model performance that are well aligned with changes that are observed in human performance as fatigue levels increase. In the next section, we describe a dual-task paradigm and a data set that has provided an opportunity to conduct an evaluation of some of the mechanisms we have proposed.

TASK CONTEXT AND RELEVANCE

Dual-tasking is a fundamental skill in naturalistic settings. Perhaps the most notorious modern example of dual-tasking is dialing and using a cell phone while driving (e.g., Salvucci, 2006), but there are countless more mundane examples, such as reading while watching TV, listening to music while typing, or checking e-mail during a meeting or presentation. Dual-tasking situations require the management of cognitive resources and the effective execution of actions in the service of each task. It is commonly observed that performance on one or both tasks degrades when they are performed together versus when each is done in isolation. In laboratory settings, dual-task paradigms provide an opportunity to examine in detail issues related to distributing cognitive resources to perform multiple tasks simultaneously. When one understands these details, a better understanding of how more complex tasks are interleaved in naturalistic contexts can emerge.

Research on dual tasking has often looked at a phenomenon referred to as the *psychological refractory period*, or PRP (Meyer & Kieras, 1997a; 1997b; Pashler, 1994; Pashler & Johnston, 1989). In a typical study, participants are presented with two tasks, each of which is a two-choice reaction time task. Critically, the stimulus for Task 2 is presented at some delay, or *stimulus onset asynchrony* (SOA), relative to Task 1. Across a variety of circumstances, reaction times for Task 2 tend to increase when the SOA is short. The PRP effect is defined as the difference between reaction time in a short-SOA condition (50 ms in the experiment described later) and the reaction time when the SOA is long (1,000 ms in the experiment).

Pashler (1994) postulated that the PRP effect reflects a central information-processing bottleneck that imposes a serial constraint on cognitive activity. However, alternative proposals suggest that the PRP effect is actually a function of strategies adopted by participants to meet the requirement that the response for Task 1 be issued first. This latter explanation has been referred to as *strategic response deferment* (Meyer & Kieras, 1997a, 1997b). Distinguishing between these alternatives has potentially important implications in complex applied tasks with time-critical components. Byrne and Anderson (2001) presented a computational cognitive model that incorporates both a cognitive bottleneck and strategic response deferment, which provided a close fit to human performance data across multiple experimental variations. This model provides the foundation for the current effort to validate a computational account of fatigue.

The data set for this evaluation comes from an experiment described in Bratzke, Rolke, Ulrich, and Peters (2007), who used a task based on the paradigm described in. In the Bratzke et al. study, Task 1 required participants to respond to an auditory signal (high versus low tone), whereas Task 2 required a response to a visual stimulus (*X* versus *O*). Task 1 responses were left-hand key presses and Task 2 responses were right-hand key presses. The three SOAs used by Bratzke et al. (2007) were 50, 200, and 1,000 ms. Participants ($n = 6$) completed 108 trials, broken into three blocks of 36, every 2 hr during 28 hr of continuous wakefulness.

The data from Bratzke et al. (2007) are presented in Figure 1 (solid black lines). Participants showed a gradual speed-up on both tasks early on, which is attributable to learning. In contrast, during the overnight hours, reaction times increased substantially for both tasks (see Figures 1a and 1b). Figure 1b shows reaction times for Task 2 as a function of the SOA. Although reaction time increases regardless of the SOA, it increases more when the SOA is shorter, which is reflected in an increase in the PRP effect (Figure 1c). The overall pattern of results is consistent with a substantial body of research investigating the interaction between sleep homeostatic processes and circadian rhythms.

The experiment and data offer an exciting opportunity to evaluate our capacity to make precise predictions regarding the effect of fatigue. Dual-task performance is an increasingly important measure as technology adds complexity and distraction to everyday activities. Simultaneously, continuous operations and activity force individuals to function and perform with inadequate sleep and at times when the circadian system is applying pressure to sleep. Unfortunately, research labs that are equipped to conduct research on fatigue—typically in a hospital setting with 24-hr staffing—often focus on physiological measures to the exclusion of detailed performance data. This limits opportunities to carefully evaluate theoretical accounts in sufficient detail. Thus, we feel that the theoretical value of the opportunity afforded by the detailed performance data reported in Bratzke et al. (2007) outweighs concerns regarding methodological details in their study, including the relatively small sample, which actually is fairly typical of sleep deprivation studies. Some limitations are discussed later, and we provide illustrations of the utility of our modeling approach as well as cautions regarding potential influences on human performance.

In discussing the degradations in performance that were observed as fatigue increased, Bratzke et al. (2007) focused on the cognitive slowing account of fatigue. Indeed, a generalized slowdown does seem to characterize the observed changes in human performance appropriately as alertness declines. In contrast to this perspective, however, we have proposed an account of fatigue based on microlapses, which are very brief gaps in cognitive processing (approximately 50 ms). The brief duration of microlapses makes them unlike the cognitive lapses mentioned earlier, which have been characterized as lasting from several to 10s of seconds (e.g., Doran et al., 2001).

A feature of our account is that microlapses can produce both slowed reactions and response lapses, depending on their probability (Gunzelmann et al., in press). In this account, microlapses are the result of decreased alertness, which interferes with the production execution process in ACT-R's procedural system. To evaluate this account in the context of dual-task performance, we applied these mechanisms directly to the model presented in Byrne and Anderson (2001). The resulting model was used to generate predictions regarding the impact of fatigue on performance in the dual-task experiment reported in Bratzke et al. (2007). The model and the fatigue mechanisms are described in the next section.

ACT-R MODEL OF DUAL-TASK PERFORMANCE

The ACT-R model of the PRP we used is nearly identical to the one presented in Byrne and Anderson (2001); only minor changes in the model's knowledge were required to accommodate the differences in stimuli between the experiments. The model contains a small number of productions, which implement a simple perceive-decide-act strategy for both tasks. Because of ACT-R's cognitive bottleneck, the decision phase of Task 2 may have to wait for

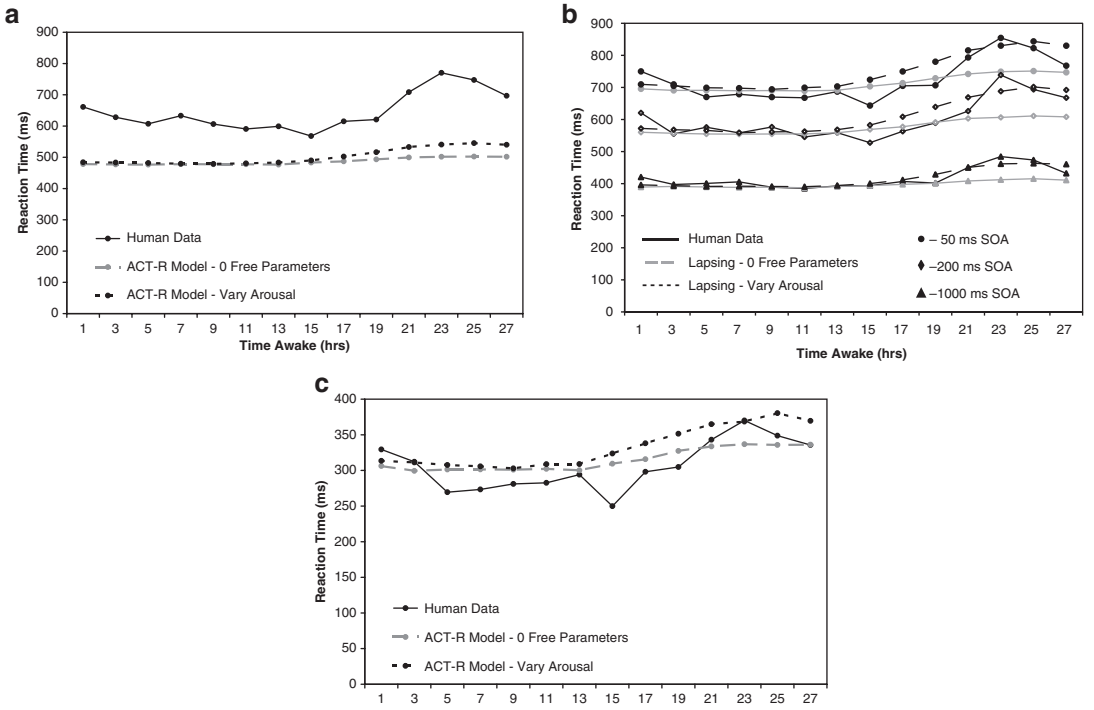

Figure 1. Human data from Bratzke, Rolke, Ulrich, and Peters (2007) and performance of the ACT-R model with and without parameter variation: (a) reaction time (ms) for Task 1, (b) reaction time (ms) for Task 2 as a function of the stimulus onset asynchrony (SOA), and (c) magnitude of the psychological refractory period effect (ms), computed by subtracting mean reaction for Task 2 at 1,000 ms SOA from the mean reaction time for Task 2 at 50 ms SOA. In all figures, human data are represented by solid black lines. Gray lines illustrate zero free parameter predictions, and short dashed lines show model performance when arousal (*G*) is allowed to vary. The human empirical data (solid lines) used in this evaluation are from "Central Slowing During the Night," by Bratzke et al., 2007, *Psychological Science*, *18*(5), p. 459, Figure 2. Copyright 2007 by Blackwell. Used with permission. All rights reserved.

Task 1 processing. Because of noise in the duration of the output stages (i.e., button presses), the model defers the response for Task 2 until the response for Task 1 has been initiated, which is done to guarantee that responses are emitted in order. This is similar to the double-bottleneck model of De Jong (1993) except that the second bottleneck is strategic (knowledge based) rather than structural (i.e., an architectural constraint), which is similar to the models proposed by Meyer and Kieras (1997a, 1997b). The first bottleneck is the serial production system, which represents an architectural constraint on central cognitive processing in ACT-R. The PRP effect occurs if either (or both) of these bottlenecks are busy when the Task 2 stimulus is presented at short SOAs, whereas the bottlenecks are free at longer SOAs. The PRP model is described in more detail in Byrne and Anderson (2001).

Mechanisms for Fatigue

The model from Byrne and Anderson (2001) provides an a priori account of human performance on the dual-task paradigm in Bratzke et al. (2007). To make predictions about the effects of fatigue, one must augment this model with mechanisms that account for degradations in cognitive performance that are associated with extended wakefulness. This involved implementing the microlapse account from Gunzelmann et al. (in press) in the PRP model from Byrne and Anderson (2001). In this

integrated model, slower reaction times, such as those observed during the overnight period in Bratzke et al. (2007), result from small numbers of microlapses that interfere with task performance in the model.

Microlapsing. The mechanisms that produce microlapses in our model were taken directly from the model in Gunzelmann et al. (in press) that performs the Psychomotor Vigilance Test (PVT), a sustained-attention task used extensively by sleep researchers investigating the impact of fatigue on cognitive performance (e.g., Dinges & Powell, 1985; Doran et al., 2001). Microlapses are implemented through a set of mechanisms that, with increasing levels of fatigue, increases the likelihood that a production execution cycle will occur in which no actions are performed. This happens when the expected utilities (U_i) of the applicable productions fail to exceed the threshold for execution, referred to as the *utility threshold* (T_u). This circumstance produces a microlapse, instantiated as a production execution cycle in which no cognitive actions are performed (lasting approximately 50 ms). Noise (ε) in the utility computation (see Equation 1) makes it possible that no production will fire on a particular cycle (because no production exceeds the threshold), followed by a cycle in which at least one production does exceed T_u and fires.

$$U_i = P_i G - C_i + \varepsilon \quad (1)$$

The frequency of microlapses in the model is a function of the *G* parameter from Equation 1 and the utility threshold (T_u). As fatigue levels increase, *G* is reduced to represent decreased alertness. Based on evidence for compensatory behaviors on the part of participants to offset the deleterious effects of sleep deprivation (e.g., Doran et al., 2001; Portas et al., 1998), we also reduced the utility threshold over time, which made it more likely that some production will fire.

The dynamics of both of these parameters have been linked to predictions of alertness from biomathematical models was used in which a linear scaling that required estimating a slope and an intercept for each mapping function. In Equation 1, P_i and C_i are production-specific parameters representing the probability and anticipated cost (in time), respectively, of achieving the goal using that production. Default values of $P_i = 1$ and $C_i = 50$ ms were used for all productions in this model.

The fatigue model also includes a mechanism that dynamically reduces arousal (*G*) during a trial. This mechanism reflects an assumption that microlapses are indicative of the process of falling asleep. Thus, *G* is decremented by a small amount each time a microlapse occurs. When a new stimulus is presented, *G* returns to its initial value at the start of each trial, which represents the arousing impact of environmental stimulation (e.g., Pilcher, Band, Odle-Dusseau, & Muth, 2007).

Gunzelmann et al. (in press) provided a more thorough description of the mechanisms and parameters as well as an evaluation of the ability of these mechanisms to capture changes in sustained-attention performance stemming from fatigue. Here we have applied these mechanisms directly to the model of the PRP from Byrne and Anderson (2001) to generate predictions about performance for the Bratzke et al. (2007) study.

MODEL EVALUATION

The direct integration of a set of mechanisms to account for fatigue into an existing model of dual-task performance offers a relatively rare opportunity in computational cognitive modeling research—the ability to make truly a priori quantitative predictions regarding performance. That is, by using CNPA to estimate alertness in the current study, and using the same linear function to map alertness predictions to ACT-R parameter values as estimated in Gunzelmann et al. (in press), we were able to generate model performance predictions with zero free parameters using the model from Byrne and Anderson (2001). The gray lines in Figure 1 present such predictions for the dual-task paradigm described in Bratzke et al. (2007), aggregated across 12,000 simulated PRP trials (approximately 111 simulated participants in the procedure used in the original study). The model captures the qualitative trends quite well, with worsening performance during the overnight hours and some recovery during the subsequent morning. However, the model is substantially faster than the humans in Task 1, and the magnitude of the model's changes in Task 2 performance clearly is not as large as in the human data.

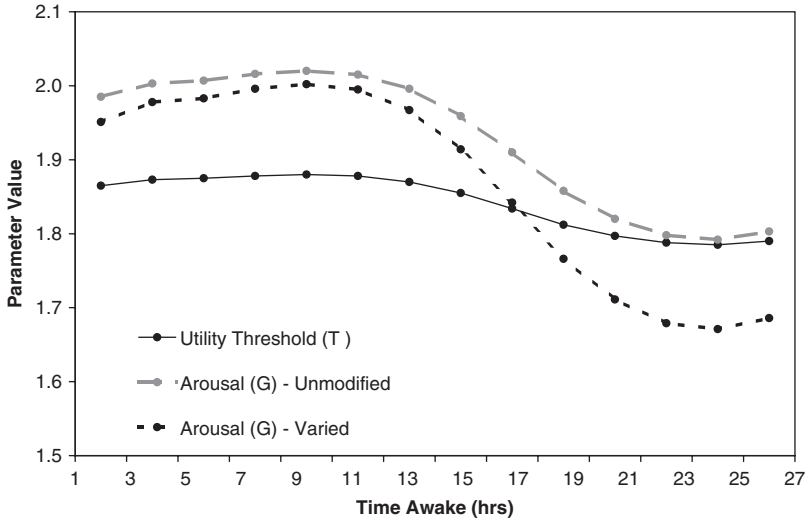

Figure 2. Values for arousal (G) and utility threshold (T_u) for the model across 28 hr of sleep deprivation. Unmodified values for *G* were estimated using the results from Gunzelmann et al. (in press). Varied values for *G* were estimated by finding the best-fitting line mapping biomathematical model predictions of alertness to *G* values for the current experiment. In both cases, T_u values were derived from Gunzelmann et al (in press).

There are a number of potential explanations for the discrepancy in the impact of fatigue, including individual differences and methodological differences related to the task and/or sleep deprivation protocol. Of these, we note that task characteristics have been found to influence motivation and performance in important ways (Caldwell & Ramspott, 1998; Drummond, Brown, Salamat, & Gillin, 2004; Pilcher et al., 2007). That is, task characteristics can influence the severity of the decrements associated with fatigue by affecting factors, such as motivation, that influence alertness. In the context of the microlapse hypothesis, this implies that different values for *G* may be necessary to capture differences in the task itself.

Figure 1 also presents data from the model when the relationship of *G* to alertness values from CNPA was allowed to vary (dashed lines; again aggregated across 12,000 trials). In this case, the model makes much more accurate predictions regarding the changes in Task 2 performance that are observed during the course of 28 hr without sleep.

Even with variations to *G*, the theoretical account is entirely consistent with the account in Gunzelmann et al. (in press) for sustained-attention performance. In addition, the dynamics of *G* across sessions are still tied to the predictions of the biomathematical model. By varying *G*, we are introducing two free parameters: the slope and the intercept of the linear equation mapping CNPA predictions of alertness to values of *G* in ACT-R.

The values of *G* using both approaches are shown in Figure 2, along with the values for T_u, which were the same in both cases. The variations of *G* relative to T_u affect the likelihood of occurrences of microlapses. In the best performance session predicted by CNPA (Session 5; following 9 hr awake), the model in which *G* was varied produced an average of 5.4 microlapses per trial, although the median (2) and the mode (0) are more appropriate, given that the distribution is highly skewed. After 25 hr awake, the average soars to 18.3, with corresponding increases in the median (10) and mode (2).

Figure 1 illustrates how the likelihood of microlapses affects overall performance. Response times increase, and the PRP effect is magnified. The modified estimates for *G* improve the model's correspondence with the data (Table 1), although the large discrepancy between the model and human data on Task 1 performance remains. This disparity appears to result from the instructions given to participants in Bratzke et al.'s (2007) study. The ACT-R model is based on task instructions that were carefully crafted to produce compliance with the PRP assumptions

TABLE 1: Comparison of the Model's Performance to the Human Data Using Root Mean Squared Deviation (RMSD) and Correlation (*r*)

	Microlapsing			
	Zero Free Parameters		Vary Arousal	
Variable	RMSD	*r*	RMSD	*r*
Task 1 average RT	168 ms	.81	149 ms	.84
Task 2 average RT (× SOA)	42 ms	.96	32 ms	.98
PRP effect	25 ms	.76	34 ms	.76

Note. RT = reaction time; PRP = psychological refractory period.

(see Schumacher et al., 1999). These instructions emphasize giving priority to Task 1 and responding to Task 1 first in addition to responding as quickly as possible to both tasks.

In contrast, Bratzke et al. (2007) instructed participants to "give the same attention to each of the two tasks" and stated that "both tasks are equally important" (Bratzke, personal communication, October 7, 2008). Such changes to the instructions can have important consequences for performance. In fact, when Pashler and Johnston (1989) instructed participants to "produce responses in a 'nice orderly rhythm'" (p. 36), Task 1 reaction times were significantly longer than when standard PRP instructions were given. Interestingly, there was little impact of the instructional manipulation on Task 2 reaction times, which appears to correspond to the current empirical data based on the contrast with the model predictions.

Our goal was not to perform a detailed evaluation of the strategy variations that are possible in the PRP. Nonetheless, it may be the case that the Byrne and Anderson (2001) model does not capture precisely the strategy used by participants in Bratzke et al. (2007) because of instructional differences. The more important question, however, is whether this discrepancy in task strategy has an impact on the validity of the conclusions made regarding the impact of fatigue. We believe that it does not. The reason is that these strategy variations involve mostly altering the precise manner in which the tasks are interleaved (Pashler & Johnston, 1989). Thus, the primary impact should be on how the two tasks are prioritized rather than on altering the fundamental processing that participants perform. Consequently, varying the model to reflect these strategic choices would not substantially alter the predictions regarding the impact of fatigue.

CONCLUSION

At a qualitative level, it is well understood that fatigue leads to worse performance in a variety of tasks. However, in any safety-critical domain in which fatigue is a factor, it is obviously useful to have more precision. Recently, biologically and physiologically inspired mathematical models have been generated that make quantitative predictions about alertness. To bridge the gap from these models to predictions about task performance, we have developed a model that instantiates alertness in a way that captures its effect on cognitive processes. We have demonstrated the predictive capacity of this model in the context of dual tasking and the PRP.

Decrements in a parameter associated with alertness in ACT-R led to microlapses, which increased reaction times for both tasks and led to larger PRP effects corresponding to observed changes in human performance. Some discrepancies in the fit of the model to the human data illustrate areas for future research, such as the particular dynamics of alertness as predicted by the biomathematical model and the potential influence of strategy on performing the dual-task paradigm used in this study. However, the progress presented in this article demonstrates how mathematical and computational process modeling can be combined to provide a more comprehensive account of human performance under the moderating influence of sleep deprivation.

The integrative approach of this research has allowed us to present a priori predictions regarding the impact of fatigue. This illustrates progress toward a major goal of research in this area, which is to make predictions about human performance in novel contexts, without the benefit of empirical data for parameter fitting. This capacity is important for both theoretical and practical, applied reasons; even approximate quantitative predictions can be crucial in applied contexts (Newell & Card, 1985). Explanatory

mechanisms that provide only descriptive accounts of known empirical data, however, will have little predictive capacity in tools and systems that are intended to reduce risk and improve performance in real-world environments. Tools based on descriptive characterizations of cognitive slowing and cognitive lapsing, for instance, are insensitive to variations in the task and are unable to generate predictions regarding the types, magnitudes, and frequencies of errors that will occur during fatigued performance. The same is true of mathematical models that quantify only relative alertness.

Without generative computational mechanisms for cognitive processing, tools can predict only relative changes in performance capacity. Similarly, cognitive architectures traditionally have ignored the moderating effects of fatigue on cognitive processing. This article provides a demonstration that our integrative approach, which leverages the strengths of both the biomathematical and the cognitive architecture approach, moves us in the direction of achieving precise, accurate predictions of future performance.

ACKNOWLEDGMENTS

The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. government. The research was sponsored by the Air Force Research Laboratory's Warfighter Readiness Research Division and by Grant 07HE01COR from the Air Force Office of Scientific Research. We greatly appreciate the willingness of Daniel Bratzke and colleagues to openly share their empirical data in support of our model development and validation efforts.

REFERENCES

Achermann, P. (2004). The two-process model of sleep regulation revisited. *Aviation, Space, and Environmental Medicine*, *74*(3), A37–A43.

Anderson, J. R. (2007). *How can the human mind occur 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*, 1036–1060.

Anderson, J. R., & Lebiere, C. L. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum.

Borbely, A. A. (1982). A two-process model of sleep regulation. *Human Neurobiology*, *1*, 195–204.

Bratzke, D., Rolke, B., Ulrich, R., & Peters, M. (2007). Central slowing during the night. *Psychological Science*, *18*, 456–461.

Byrne, M. D., & Anderson, J. R. (2001). Serial modules in parallel: The psychological refractory period and perfect time sharing. *Psychological Review*, *108*, 847–869.

Caldwell, J. A. (2003, Fall). Wake up to the importance of sleep for air safety! *Flightline*, pp. 30–33.

Caldwell, J. A., & Ramspott, S. (1998). Effects of task duration on sensitivity to sleep deprivation using the Multi-Attribute Task Battery. *Behavior Research Methods, Instruments, and Computers*, *30*, 651–660.

De Jong, R. (1993). Multiple bottlenecks in overlapping task performance. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 965–980.

Dean, D. A., III, Fletcher, A., Hursh, S. R., & Klerman, E. B. (2007). Developing mathematical models of neurobehavioral performance for the "real world." *Journal of Biological Rhythms*, *22*, 246–258.

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

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. 97–128). Chichester, UK: Wiley.

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, and Computers*, *17*, 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*, 253–267.

Drummond, S. P. A., Brown, G. G., Salamat, J. S., & Gillin, J. C. (2004). Increasing task difficulty facilitates the cerebral compensatory response to total sleep deprivation. *Sleep*, *27*, 445–451.

Gallup. (2002). *National survey of distracted and drowsy driving attitudes and behaviors: 2002. Volume 1: Findings report*. Washington, DC: National Highway Traffic Safety Administration. Retrieved January 29, 2009, from http://www.nhtsa.dot.gov/people/injury/drowsy_driving1/survey-distractive03/index.htm

Gunzelmann, G., Gluck, K. A., Kershner, J., Van Dongen, H. P. A., & Dinges, D. F. (2007). Understanding decrements in knowledge access resulting from increased fatigue. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the twenty-ninth annual meeting of the Cognitive Science Society* (pp. 329–334). Mahwah, NJ: Lawrence Erlbaum.

Gunzelmann, G., Gross, J. B., Gluck, K. A., & Dinges, D. F. (in press). A computational account of changes in psychomotor vigilance as a consequence of fatigue. *Cognitive Science, 33*.

Heuer, H., Kleinsorge, T., Klein, W., & Kohlisch, O. (2004). Total sleep deprivation increases the costs of shifting between simple cognitive tasks. *Acta Psychologica*, *117*, 29–64.

Horne, J., & Reyner, L. (1999). Vehicle accidents related to sleep: A review. *Occupational and Environmental Medicine*, *56*, 289–294.

Hursh, S. R., Raslear, T. G., Kaye, A. S., & Fanzone, J. F., Jr. (2006). *Validation and calibration of a fatigue assessment tool for railroad work schedules: Summary report* (Report DOT/FRA/ORD-06/21). Washington, DC: U.S. Department of Transportation, Office of Research and Development.

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

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). *The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data* (Tech. Report DOT HS 810 594). Washington, DC: National Highway Traffic Safety Administration.

Klerman, E. B., & St. Hilaire, M. (2007). On mathematical modeling of circadian rhythms, performance, and alertness. *Journal of Biological Rhythms*, *22*(2), 91–102.

Mallis, M., Mejdal, S., Nguyen, T., & Dinges, D. (2004). Summary of the key features of seven biomathematical models of human fatigue and performance. *Aviation, Space, and Environmental Medicine*, *75*(3), A4–A14.

Meyer, D. E., & Kieras, D. E. (1997a). A computational theory of executive control processes and multiple-task performance: I. Basic mechanisms. *Psychological Review*, *104*, 3–65.

Meyer, D. E., & Kieras, D. E. (1997b). A computational theory of executive control processes and multiple-task performance: II. Accounts of psychological refractory period phenomena. *Psychological Review*, *104*, 749–791.

Mitler, M. M., Carskadon, M. A., Czeisler, C. A., Dement, W. C., Dinges, D. F., & Graeber, R. C. (1988). Catastrophes, sleep, and public policy: Consensus report. *Sleep*, *11*(1), 100–109.

National Transportation Safety Board. (1995). *Factors that affect fatigue in heavy truck accidents: Vol. 1* (NTSB Safety Study Report No. NTSB//SS-95/01 [PB95-9170012]). Washington, DC: Author.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Newell, A., & Card, S. K. (1985). The prospects for psychological science in human–computer interaction. *Human–Computer Interaction*, *1*, 209–42.

Pack, A. I., Pack, A. M., Rodgman, E., Cucchiara, A., Dinges, D. F., & Schwab, C. W. (1995). Characteristics of crashes attributed to the driver having fallen asleep. *Accident Analysis and Prevention*, *27*, 769–775.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*, 220–244.

Pashler, H., & Johnston, J. C. (1989). Chronometric evidence for central postponement in temporally overlapping tasks. *Quarterly Journal of Experimental Psychology*, *41A*, 19–45.

Pilcher, J. J., Band, D., Odle-Dusseau, H. N., & Muth, E. R. (2007). Human performance under sustained operations and acute sleep deprivation conditions: Toward a model of controlled attention. *Aviation, Space, and Environmental Medicine*, *78*(5-II Suppl.), B15–B24.

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*, 8979–8989.

Salvucci, D. D. (2006). Modeling driver behavior in a cognitive architecture. *Human Factors*, *48*, 362–380.

Schumacher, E. H., Lauber, E. J., Glass, J. M., Zurbriggen, E. L., Gmeindl, L., Kieras, D. E., et al. (1999). Concurrent response-selection processes in dual-task performance: Evidence for adaptive executive control of task scheduling. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 791–814.

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.

Glenn Gunzelmann is a senior research psychologist at the Air Force Research Laboratory in Mesa, Arizona. He received his PhD in cognitive psychology from Carnegie Mellon University in Pittsburgh, Pennsylvania, in 2003.

Michael D. Byrne is an associate professor in the Department of Psychology at Rice University in Houston, Texas. He received his PhD in experimental psychology from Georgia Institute of Technology in Atlanta in 1996.

Kevin A. Gluck is a senior research psychologist at the Air Force Research Laboratory in Mesa, Arizona. He received his PhD in cognitive psychology from Carnegie Mellon University in Pittsburgh, Pennsylvania, in 1999.

L. Richard Moore Jr. is a software engineer at Lockheed Martin in Mesa, Arizona. He received his MS in applied psychology from Arizona State University in Tempe in 2008.

Date received: March 14, 2008
Date accepted: February 11, 2009