Research Requirements for Operational Decision-Making Using Models of Fatigue and Performance

Karl E. Friedl, Melissa M. Mallis, Stephen T. Ahlers,
Stephen M. Popkin, and Willard Larkin

Friedl KE, Mallis MM, Ahlers ST, Popkin SM, Larkin W. ***Research requirements for operational decision-making using models of fatigue and performance.*** **Aviat Space Environ Med 2004; 75(3, Suppl.):A192–9.**

Sustained human performance is critical to job and mission success in many federal agencies including national defense, aerospace exploration, and transportation. For the responsible agencies, applications of the basic biomedical and applied human factors science provide the best available solutions to help individuals perform more effectively and with increased safety. Key products of this research are biomathematical models that predict periods of impaired performance, with applications in planning tools, real time monitoring, and intervention decision aids. Since it is difficult to quantify the number of judgment errors or accidents averted, metrics of success for fatigue management systems must be largely based on the accuracy of performance predictions derived from laboratory-based research studies and the extent to which such results can be generalized to the field environment. Performance metrics must, at a minimum, be correlated with occupational task performance to demonstrate relevance to real-world applications. This paper outlines broad goals for human effectiveness research related to fatigue, alertness, and performance. To advance from the present state of knowledge to useful predictive models requires a well-coordinated commitment from federal agencies. Users should be made aware that current models and tests are not likely to encompass all of the aspects of human performance that are relevant to field environments and occupations.

Keywords: biological models, task performance and analysis, personnel staffing and scheduling, psychomotor performance, cognition, physiology, neuropsychological tests, sleep deprivation, fatigue, physiopathology, biomathematical models, alertness, performance.

CURRENT TECHNOLOGICAL advances have resulted in the development of vehicles that can operate continuously for hundreds of maintenance-free hours, with impending failure predicted from embedded sensors and performance status monitored automatically from sensor data, system history, and predictive models. Prognostics and diagnostics of comparable accuracy and specificity do not yet exist for the human engaged in sustained performance or working on rotating shifts. Nevertheless, fatigue and performance models can potentially provide an important advantage when human performance is critical.

The ability to predict performance levels based on recent sleep/wake history and other factors such as emotional, cognitive, and physical loads will facilitate schedule planning and realistic risk management. In combination with physiological inputs reflecting moment-to-moment fluctuations in alertness and performance levels, it should be possible to provide real-time information on safe and effective performance capability. Depending on the sensor information provided (e.g., physiological monitoring and embedded performance measures), prescriptive models, ideally, should be able to combine real-time information with recent sleep/wake history and circadian phase information to suggest courses of action to sustain alertness, performance, and safety (e.g., interventions such as optimally timed naps or doses of caffeinated products). Additionally, the model should be able to recommend alternate sleep/wake schedules to identify periods of compromised performance. Development of such a fatigue management tool is still in the very earliest stages. Until these models are developed, providing a tool to decrease the relative risk of human errors and failure, this significant gap in technology leaves the human the weak link in human-machine interactions and coordination. This paper outlines general research needs for fatigue management tools without any intention to limit or dictate paths of novel scientific exploration that may lead to revolutionary breakthroughs.

Agency Requirements

Human performance is a key concern for several federal agencies, notably the Department of Defense (DOD), the National Aeronautics and Space Administration (NASA), and the Department of Transportation (DOT). Biomedical research in these agencies is generally more focused on promotion of human safety and effectiveness rather than enhancement of health and disease prevention. For all human-machine systems, information processing is critical. Although information technology can mitigate some aspects of information overload, it cannot eliminate the need for alert decision-making by a human operator. In the Army's plans for the Future Force, hundreds of reports from remote sensors and unmanned aerial vehicles will be flowing into command centers where human attention and expertise must be available. The U.S. Air Force also deploys a

---

From the U.S. Army Medical Research and Materiel Command, Frederick, MD (K. E. Friedl), the NASA-AMES Research Center, Moffett Field, CA (M. M. Mallis), the Office of Naval Research, Arlington, VA (S. T. Ahlers), the Department of Transportation, Volpe Center, Cambridge, MA (S. M. Popkin), and the Air Force Office of Scientific Research, Arlington, VA (W. Larkin).

The first author, Karl Friedl, is now Commander, U.S. Army Research Institute of Environmental Medicine, Natick, MA.

Address reprint requests to: Karl Friedl, P.O. Box 29, Natick, MA 01760; karl.friedl@na.amedd.army.mil.

Reprint & Copyright © by Aerospace Medical Association, Alexandria, VA.

variety of unmanned air vehicles, often many at the same time, but flight control and situational awareness remain a human responsibility. Chronobiological considerations, as well as workload, will continue to be critical to Air Force missions, with most airmen involved in command and control tasks that continue around the clock and in every time zone around the globe. For the U.S. Navy and Marine Corps, sustaining forces in any region of the world under conditions of increased pace of deployments places significant demands on every individual operator. Seamen are being cross-trained and used in a wider variety of functions with greater levels of responsibility, as ships' crews are being reduced in size by as much as 75%. An increased understanding of individual variability is critical to development of useful performance prediction models and necessary for real time status assessment and near term performance predictions for individual operators.

NASA has special concerns about individual human reliability, mental status monitoring, and optimal performance status for long-range missions such as the Mars mission. The development of biomathematical models of performance and alertness will contribute to the knowledge that will allow for future long-term space missions. Specifically, the development of such models will contribute to two specific agency strategies to accomplish the NASA mission: 1) enhance core scientific capabilities and processes to ensure safety and mission success and reduce cost; and 2) ensure that all NASA work environments, on Earth and in space, are safe, healthy, environmentally sound, and secure. There is a particular focus in the development of optimal sleep/wake and work schedules to increase efficiency and safety in aerospace operations. For example, the 90-min orbital cycle (day/night) of the space shuttle, combined with a rest/activity cycle that is frequently shorter, and sometimes longer, than 24 h, have the potential to produce a misalignment between the phase of the circadian pacemaker and the sleep/wake cycle, leading to circadian desynchronization and performance degradation. Development of optimal scheduling strategies and countermeasure application in aerospace operations will minimize the adverse effects of extended operations and contribute to NASA's goal of optimizing crew safety, well being, and performance.

The DOT has responsibility for the largest number and widest variety of continuous performance and/or shift workers, including long haul truckers, train engineers, railroad and pipeline dispatchers, airplane pilots, air traffic controllers, and those involved in maritime activities. In the DOT, fatigue management is now viewed as an important adjunct to hours of service regulations that were instituted to meet the opposing demands of productivity and safety in the nation's transportation system. All of the agencies have interests in fostering appropriate planning and scheduling of their workforce, providing real-time physiological and fatigue monitoring capabilities, and developing highly reliable fatigue management tools.

Examples of some of the applications of performance models are listed below:

- Work and rest scheduling
- Safety policies
- Hours of service regulations
- Ultra-long-range schedules and long-haul flights
- Long duration space missions
- High tempo operations
- Leader effectiveness scheduling
- Personal assistant/planner
- Education
- Accident investigation
- Collaborative decision aids (decision assist for fatigue countermeasures)
- Mission rehearsal tools (allow leaders to learn their limits)
- Improved situational awareness (real-time prediction of performance status)

Planning to accommodate typical human tolerances for safety and effectiveness could include both manning levels and task design. Reliable information on the effects and effectiveness of various potential interventions may allow the generation and comparison of alternate solutions or development of emergency steps and countermeasures. In the DOT, such information could be used to create benchmarks by which an agency could evaluate a company's unique fatigue management plan rather than provide prescriptive guidance. More reliable predictions could limit or warn about worker schedules based on current fatigue status (perhaps with physiological monitoring) and other required inputs. As they improve in reliability, fatigue models could be used to guide policies and assist decisions for sleep periods or alertness countermeasures, eventually accounting for multiple influences such as recent diet, stimulant use, physical and psychological loads, and other relevant factors for the individual. Sophisticated modeling will predict behavior in complex jobs, including those with crew interactions and the need to synthesize knowledge from many disparate pieces of information. Ultimately, the models should be adaptive, with capabilities to recognize prediction deficiencies and automatically seek out better data as well as provide statistically based predictions that quantify the level of uncertainty.

Characteristics of a Useful Model

In order to be useful in an operational environment, a model must be validated in the specific operational field settings and be proven to be useful and useable for prediction of some important aspect of performance or related outcomes. The characteristics of fatigue performance models are listed here.

Minimum requirements are:

- Cognitive performance based
- As simple and utilitarian as possible with respect to algorithm and input requirements
- Proper validation in lab and field
- Defined statistical properties
- Known sensitivity to real-world data
- Predicts effects of simple countermeasures (naps, stimulants) and interactions
- Predicts performance recovery
- Common data formats and protocols for data input

- Common output and hooks that may be used as input by other software

Desirable prediction capability improvements should be able to incorporate the following:

- Effects of increased cognitive workload and demand
- Effects of other psychological stress (i.e., anxiety)
- Effects of transmeridian travel
- Effects of work schedule and sleep schedule irregularity
- Role of light exposure
- Individual differences
- Effects of fatigue on team performance
- Effects of environmental stressors (e.g., exercise, temperature, noise)

Before operational testing, a model should be rigorously developed in the laboratory (10). If a model is not first validated in a controlled laboratory research setting, where independent variables can be tightly controlled, there is the risk of wasting research and tax dollars on research validation in an operational environment without the specific identification of the operational variables of interest. The model must provide confidence intervals on performance predictions. A phrase such as "68% of baseline performance" has little meaning without knowledge of the individual variability and the accepted performance levels in the specific environment. Appropriate operationally relevant inputs to the model must be defined and available. For example, different levels of vigilance might be required for the landing phase of an aerospace operation than are required during a submarine watch. Only after these acceptable operational performance levels are defined will the output of the model be useful in the operational environment. Also, a person's sleep/wake history alone may be insufficient, since some tasks are relatively insensitive to even severe sleep deprivation, and some individuals are relatively impervious to the effects of sleep loss. Performance predictions may be obvious at the extremes of well rested and extremely sleep-deprived, where nearly all humans will perform optimally or be impaired regardless of fatigue countermeasures employed; the area of modeling interest is between these extremes, where individual variability and other factors that moderate fatigue effects are important in the model. The effects of basic restorative naps or recovery sleep, the use of common stimulants such caffeine, and the interactions of these fatigue interventions should be reasonably predicted by any basic model. Later refinements of a basic model should target some of the planned product improvements listed at the bottom of the above list.

The difficulties facing behavioral researchers are compounded by a prevalent perception that human behavior has been well described over human history and new research is only likely to provide trivial refinements of what is already known. Likewise, sleep and alertness physiology is inherently interesting to the general public and everyone is an "expert" on the basis of personal experience. This perception that fatigue and performance relationships are well understood makes it even more important for scientists to provide only well-validated models for prediction of relevant operational performance and to carefully manage user expectations by clearly presenting guidance on the appropriate use of the models and their limitations.

Prediction Outcomes: Alertness, Performance, or Safety

Several formidable barriers impede progress toward implementation of a useful fatigue management model, but the overarching challenges are in the outcome measures. First is the lack of a common, well-defined, and accepted nomenclature for discussing relevant concepts such as acceptable performance, alertness, and fatigue levels. Even terms such as "fatigue" and "napping" need definition. Such definitions must be established up front, otherwise ambiguity as to what the models are predicting and the criteria by which they are evaluated will persist. Next, the specific outcome measures against which models were developed need to be explicitly defined and related to operational performance. While the original two-process model that underlies most of the current models was based on measures of slow-wave sleep (4), models have perhaps too casually migrated to the prediction of laboratory measures, such as measures of throughput, reaction time, and sustained attention, and are then applied to predict overall job performance, which may require very different aspects of performance than was required by the laboratory tests. Models are needed to predict complex task performance that may involve multiple aspects of cognitive function, such as decision making and psychomotor control in a sentry engaged in prolonged vigilance duties; multiple physical and psychological challenges such as thermal strains or physical demands and imminent threats that may activate or distract; and team efforts in which unit cohesion, communications, and other aspects of performance are affected by fatigue. There is high risk in using any of the currently available models to predict individual effectiveness or particular high reliance schedule demands, with all the uncertainties in even the definition of relevant performance outcomes that may not match the intended application. These challenges highlight some of the most urgent research requirements, including: 1) the need for useful and valid indicators of relevant operational performance, 2) the need to understand the basis of individual differences in performance capacity during both total and partial sleep loss, and 3) better characterization of the relationship between measures of human alertness and mission success.

Simple and convenient tests for research data collection such as the Psychomotor Vigilance Task need to be further validated against operationally relevant performance measures to characterize and delimit the generalizability of the Psychomotor Vigilance Task or other research test-based predictions (12). Development of effective measures, especially noninvasive tests that can be used transparently during actual performance-demanding tasks, may provide new research insights that cannot be achieved with testing that requires interruption of performance routines. These same tests may also provide even more effective predictions with real time inputs on the performance status of the individual. Less obtrusive physiological measures, such as occulomotor responses (21) or frontal lobe blood flow measured by near infrared spectroscopy (25), may provide additional indicators of change in individual performance, but physiological predictors of behavior are far less mature than simple performance tests. Neuropsychological testing such as the Automated Neuropsychological Assessment Metric may dissect out specific aspects of impaired performance, but have not yet been adequately related to tasks involving vigilance, high workloads, or substantial complexity (13) during operational demands.

Another approach is to experimentally measure performance in a complex synthetic task environment (STE). Significant advances in this field have been pioneered by the Air Force, with quantifiable models of human performance ("human processing units") for a command and control center and also for flight control of unmanned aerial vehicles (23). There are several advantages to STEs over simpler tests. They reproduce the cognitive elements of a complex task, as well as some of the essential operational components. Performance on an STE can be validated and benchmarked against real operational tasks. In the case of the unmanned aerial vehicle tasks, performance on the STE has been validated against the real tasks with experienced operators and has even been adopted as a trainer by some Air Force operators because of the high fidelity of the cognitive aspects. The definition of "good," "better," and "best" levels of human performance can be determined, and quantified, using the performance of real-world experts. Also, once a synthetic task has been created, models for human performance in the task, or for "optimal" performance as well, can be written. In the case of the unmanned aerial vehicles, several such models have been written, including at least one in the production-rule computational language of ACT-R. With such a computational model, the important parameters of the model can be explored. For example, manipulating operator fatigue could test a hypothesis on how operator fatigue affects an outcome such as working memory. Running sessions with fatigued individuals to test the hypothesis might also produce an alternate model suitably adjusted to predict human performance while fatigued, or perhaps fatigue could be scaled against other factors such as the comparative importance of noise in the headphones. This STE approach permits efficient and controlled testing that cannot be achieved as readily in actual operational environments. Production-rule algorithms and other nontraditional approaches to the study of human performance may provide new insights into quantification of human performance and the effects of sleep/wake dynamics on performance.

Need for Basic Research

Research in the federal portfolio moves along two convergent tracks: basic research into the biology of behavior that at any time can produce new knowledge for revolutionary advances, and applications of the science that provide best available solutions to current problems. These two tracks are managed quite differently, with basic research investments focused on creating an environment for discovery and acquisition of knowledge, and applied research focused on the best near-term solutions to an operational problem. Inquiries into why we sleep and the physiological basis of behavior will advance theoretical and experimental efforts and expand the possible, providing a strong scientific basis for future fatigue and performance models. Technological barriers and critical technologies for basic research investment include:

- Human performance markers and test methodologies
- Behavioral modeling and bioinfomatics/chaos theory tools
- Circadian physiology
- Genomics and proteomics of sleep resilience variability (individual differences)
- Neural circuits and plasticity
- Functional neuroimaging
- Commonly agreed-upon terminology

There needs to be a continued investment in the development of physiological markers linking brain and behavior, such as the initiatives in functional neuroimaging and behavior, and regional brain metabolic physiology (blood flow, oxygen and glucose uptake, electrical activity, localized biochemical activity measured by magnetic resonance spectroscopy, etc.) (2,11,22). Neuropsychological test batteries such as the Automated Neurophysical Assessment Metric (ANAM) are largely computerized versions of paper and pencil tests; even more interesting (and perhaps relevant) would be tests that take advantage of current computing power and gaming applications to probe specific, higher order brain functions. For example, visual illusions that provoke disorientation might provide a powerful challenge test that is highly sensitive to fatigue and reflects an important aspect of performance capability. The mathematical modeling itself requires basic research investment to foster development of new approaches to capabilities ranging from bioinfomatics to the handling of the complexities of human behavior, including both social and environmental interactions, as well as determining the nature of relationships between behavioral and physiological variables. One novel approach is detailed in this journal issue by Reifman (20), who described the use of hybrid models that combine neural networks with prior process knowledge. Recent published studies suggest that there are populations of individuals whose need for sleep varies significantly (1); genomics and proteomics of sleep and fatigue resilience may address some of this variability with elaboration of heritable differences as well as an understanding of mechanisms and regulators that could in turn provide useful biomarkers and interventions in fatigue management (8). Basic research investments in Parkinson's disease by the Army are providing a broader understanding of fundamental mechanisms that may explain common pathways of stimulants such as caffeine on alertness and cognition via adenosine receptors and intracellular biochemical cascades (14,17). Such re-

TABLE I. TECHNOLOGY FORECAST FOR FATIGUE MANAGEMENT MODELS.

	Current baseline	Near term (2006)	Midterm (2010)	Far term (2020)
Product	Automated "subject matter expert" guidance	Mission/work schedule planning and evaluation tool	Predictive model	Adaptive model
Description	Compendium of best available subject matter expertise in computerized format	Group level performance predictions about effects of schedules and simple interventions and interactions	Predictions of individual performance based on sleep, sympathetic activation, and other factors; more complex interactions predicted	Comprehensive predictions of relevant performance and effects of various courses of action
Applications	Expert guidance on work/mission schedules	Accurate estimates of performance in specific scenarios	Individual performance predictions for customized planning	Real-time status monitoring and decision assist programs
Characteristics	Collection of rules	Mathematical stochastic model related to valid and relevant task performance	Mathematic model dependent on input of personal history factors	Advanced bioinfomatics technology with intelligent and adaptive capabilities
Research Needs		Data on effects and interactions of fatigue interventions	Better understanding of individual variability in fatigue resilience	New methods of biomathematical modeling

search may provide a sound basis for models of the interactions of stimulants and other factors affecting human performance. Much more needs to be explored in understanding the circadian "process" portion of the model, especially in terms of how it is modified by activation, emotional stress, other environmental stressors, and light (9,15); neuroplastic aspects that explain adaptive changes; and neurochemical regulators and markers associated with sleep and circadian processes (7,16). Ultimately, with more basic research, human behavior simulators might be constructed from this knowledge, but for now, the effort is clearly focused on modeling and not on simulation.

Stages of Development

Fatigue and performance modeling is likely to follow from a rough collection of rules based on limited data, to reliable probabilistic predictions of group performance, to a tool capable of high fidelity prediction of individual performance (**Table I**). This calls for a combination of efforts, ranging from mathematical modeling of existing data to new psychophysiological experiments.

Current baseline: The current fatigue models are useful in providing what amounts to the equivalent of "subject matter expertise" for optimization of general mission and work schedule planning. However, the extent to which these model predictions are reliable, much less generalizable, has yet to be determined. The primary evaluation metric is a traceable provenance of the incorporated knowledge (e.g., annotated source book). This type of model can only generally describe current fatigue state (e.g., extent of sleep debt) and make broad generalizations about what that probably means for performance. The application is thus an all-in-one mission or work/rest schedule planner. At best, this might provide rough categories of risk, when performance would be expected to be: reduced for sustained vigilance or other tasks requiring concentration, markedly reduced with increased judgment errors, and at great risk of frank sleep onset, requiring continuous monitoring and substantial interventions to alertness. Despite the limitations, this first stage of development is immediately useful as a planning tool, essentially summarizing the current knowledge into a best estimate of how to plan rest periods and other activities, given the specific requirements of a mission. Without this, the alternative is that sleep/wake history, circadian, and other factors might otherwise be given little consideration in a planning process.

Near-term models: With greater understanding of sleep and alertness processes that will come from better definition of outcome measures (including the linkage between research measures and real-world outcomes), data from new studies using currently available technologies, and teaming with mathematical modelers, the first truly useful predictive models could be achieved in short order. While these models should provide new and better predictive capabilities, they will also be hypothesis-generating, identifying gaps in our understanding to direct new experiments. Quantitative outputs must include quantitative estimates of the confidence intervals for group mean predictions. Labor intensive, recent "sleep dose" studies (3,26) and the effects of breaks (19) are vitally important to ensuring development of valid models based on data. These models also call for much more biomedical data to account for interactions of factors that will modify performance. For example, Buguet and his colleagues (5) have assembled a conceptual framework to describe the observed effects in multiple studies of environmental factors (e.g., exercise, ambient heat and cold) on slow-wave and other aspects of sleep architecture. Even postural changes can affect EEG and performance in sleep-deprived individuals (6). Follow-on studies to translate environmental effects on sleep architecture to changes in subsequent performance are urgently needed (24). Fatigue effects on performance also need to be tested in a realistically noisy environment (18), perhaps using synthetic task environments.

Midterm models: Later stages of model development should provide predictions for pertinent groups and provide statistical bounds for the information. This would include accurate predictions of fundamental perturbations including changes in circadian phase through light and other zeitgebers, especially with travel across time zones and variable work/rest schedules; the influence of naps and stimulants; effects of emotional, cognitive, and physical load; and the influence of basic environmental factors such as heat, cold, and hypoxia. At the group level, it would also consider various types of performance effects, and how key aspects of performance vary across the day and with fatigue (i.e., resulting from extended "time on task"). Prediction capabilities should take into account major factors influencing variability that improve the generalized model for individual prediction. This might include information from genetic markers such as single nucleotide polymorphisms, challenge tests with sleep deprivation and time of day performance, and recent emotional/physical load and restorative sleep history. All of the agencies have interests in providing real time monitoring capabilities that may include sleep actigraphy along with other physiological sensors and embedded performance assessments providing accurate estimates of an individual's alertness status and predicted performance capability. This modeling effort calls for a high level of sophistication, turning databased predictions into knowledge, such as explaining what predictions about performance mean to risk of injury; vigilance capabilities; discrimination and judgment, etc.; how these changes affect plans; and providing courses of action in a decision assist mode. Well-validated models may be used to assess an individual's effectiveness and safety based on predictions from recent history and current status measurements.

Far-term models: A more sophisticated phase of modeling awaits advances in information technologies and other breakthroughs, including in our understanding of sleep and how it restores and sustains cognitive performance and alertness. This could lead to adaptive, self-assessing systems that can improve the available information by identifying weaknesses of the predictions and the precise data elements that will improve accuracy.

The Vision: What Could Be

Long range goals can serve to crystallize research programs, even if the precise path to those goals is not clear. Thus, one can easily imagine the future benefits of today's federal investment in fatigue and performance research and modeling, benefits that will include enhanced productivity and safety in virtually all operational environments.

Consider, for example, the following scenario. In the year 2025, an international emergency relief mission is deployed from the U.S. to the Indonesian Islands following a major volcanic eruption and tsunami. While the plane is being loaded and prepared for the mission, the pilots are prepared for the journey with a prescriptive program of activity, sleep, diet, and a bright light exposure regimen that helps ensure that alertness and cognitive capabilities are optimized at the appropriate times. Taking into consideration the individual sleep histories and circadian rhythms of each pilot (data that was automatically and telemetrically downloaded from their wrist-worn actigraphs as they entered the cockpit), as well as current mental status, the sophisticated model in the onboard computer provides individualized recommendations. These might include in-flight timing of meals, rest, naps, caffeine and melatonin dosing, and light exposure to optimize pilot performance during critical periods of flight (e.g., takeoffs and landings) and minimize the effect of rapid travel across multiple time zones, with optimization of the adjustment to the new light/dark cycle at the final destination. In addition, electromagnets in a special cap worn by the pilots are activated at optimal times and at specific frequencies to stimulate the retinohypothalamic pathway, hypothalamus, and thalamus to facilitate alertness and resynchronization of circadian rhythms to optimize the wake and sleep times of the two-pilot crew in a complimentary manner (i.e., so that at least one pilot is always sufficiently alert and cognitively prepared to deal with any in-flight emergencies that may arise).

Although the on-board fatigue management program predicts each pilot's alertness and performance capacity to within 95% specificity, it is still not perfect (i.e., error variance remains), so physiological status monitoring feedback systems are onboard to provide backup safety and to periodically recalibrate the system for each individual pilot. Thus, for example, the same special cap that provides electromagnetic stimulation also contains miniaturized near-infrared sensors to monitor changes in blood flow in the frontal lobes, varying as a function of alertness and cognitive load. Therefore, if several hours into the trip, a slight drift in pilot concentration is detected by these and other sensors (e.g., EEG signals, ocular movements, and slow eyelid closure, each of which would provide a portion of the convergent evidence indicating impaired alertness), the system would automatically trigger alerting stimuli (e.g., a jet of cool air on the nape of the drowsy pilot's neck, release of an alarm aroma of peppermint from a compartment on the microphone, or 5-decibel incremental increases in radio volume). The fatigue management program would also be informed of this lapse in alertness, and would use the information to recalibrate the performance predictions for that pilot, as well as provide recommendations for further interventions (e.g., chewing a 100 mg caffeinated gumball for 5 min).

Continued or subsequent lapses in attention would trigger presentation of more intrusive stimuli such as a message projected on the retina indicating that it is imperative that other validated fatigue countermeasures be initiated to ensure safe cockpit performance. It might recommend an earlier-than-scheduled switch with the other pilot, who, on being unexpectedly awakened from a nap containing slow-wave sleep, would chew a 100 mg caffeine gumball to overcome sleep inertia effects before taking the controls. For more critical lapses at critical points in the flight when time available to initiate interventions is strictly limited, an electrically activated muscle contraction to the side of the face and a child's voice urging the adult to "wake up, daddy!" would be used to trigger autonomic emergency responses. The release of adrenaline that results from presentation of these stimuli and from the sudden realization of impending danger would sustain alertness and performance for the 90 s required to stabilize the situation before initiating other, longer-acting fatigue countermeasures.

However, in this story, because the pilots followed their individualized fatigue management model-prescribed regimens of naps, meals, bright light exposure, and caffeine intake, the onboard physiological monitors detected no unscheduled instances of impaired alertness or performance, and the flight was accomplished with no problems. The crew arrives safely in Borneo, primed for a restorative sleep from which they will awake synchronized to local time. For this they could thank the fatigue management technologies that were developed for soldiers in continuous operations, transcontinental pilots, long haul truckers, astronauts preparing for aerospace operations such as the Mars mission, rail workers on long and rotating shifts, and submariners deprived of external time cues; all of which was ultimately made possible by a coordinated research effort to develop valid predictive models that adequately describe fatigue and performance.

CONCLUSIONS

Although many leaders profess that human effectiveness is critical to mission success, resources tend to be committed to tangible items instead of tools to understand, predict, and optimize conditions for human behavior. There is a common assumption that this is all well known and simply needs to be "looked up" somewhere. This easy dismissal of the need for more research is further reinforced by the ready willingness of entrepreneurs to cobble together a best available solution for customers without much concern for validating their tool. "If a model already exists and is in use, why do we need further work in this area?" Currently, there is no regulation of these models to protect users against real harm that could come through inappropriate use. This calls for extra caution on the part of developers to ensure scrupulous review and validation of any model applications. It is likely that federal agencies will develop a validation process that would provide legitimate applications; for example, these officially validated models could be used for work schedule development, perhaps even in place of current hours of service legislation.

Although generally undervalued, this area of brain and behavior research is on the critical path for each of several federal agencies. Modelers and sleep researchers must team to explore the most compelling expression of the known relationships. Predictions from these efforts will guide the studies to test the premises and to fill the gaps through subsequent iterations of development.

ACKNOWLEDGMENTS

The substantive review and contribution to the development of the sustained operations vignette by Dr. Tom Balkin is gratefully acknowledged, as are the helpful suggestions of Drs. Roy Vigneulle and Stephen Grate.

The views and opinions contained in this report are those of the authors and should not be construed as an official position, policy, or decision of any federal agency unless designated by official documentation.

REFERENCES

1. Aeschbach D, Sher L, Pstolache TT, et al. A longer biological night in long sleepers than in short sleepers. J Clin Endocrinol Metab 2003; 88:26–30.
2. Balkin TJ, Braun AR, Wesensten NJ, et al. The process of awakening: A PET study of regional brain activity patterns mediating the re-establishment of alertness and consciousness. Brain 2002; 125:2308–19.
3. Belenky G, Wesensten NJ, Thorne DR, et al. Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: A sleep dose study. J Sleep Res 2003; 12:1–12.
4. Borbely AA. A two process model of sleep regulation. Hum Neurobiol 1982; 1:195–204.
5. Buguet A, Cespuglio R, Radomski MW. Sleep and stress in man: An approach through exercise and exposure to extreme environments. Can J Physiol Pharmacol 1998; 76:553–61.
6. Caldwell JA, Prazinko B, Caldwell JL. Body posture affects electroencephalographic activity and psychomotor vigilance task performance in sleep-deprived subjects. Clin Neurophysiol 2003; 114:23–31.
7. Chapotot F, Buguet A, Gronfier C, Brandenberger G. Hypothalamo-pituitary-adrenal axis activity is related to the level of central arousal: Effect of sleep deprivation on the association of high-frequency waking electroencephalogram with cortisol release. Neuroendocrinol 2001; 73:312–21.
8. Cirelli C. How sleep deprivation affects gene expression in the brain: A review of recent findings. J Appl Physiol 2002; 92:394–400.
9. Dijk DJ, Neri DF, Wyatt JK, et al. Sleep, performance, circadian rhythms, and light-dark cycles during two space shuttle flights. Am J Physiol 2001; 281:R1647–64.
10. Dinges DF. Critical research issues in development of biomathematical models of fatigue and performance. Aviat Space Environ Med 2004; 75(3, Suppl.):A181–91.
11. Hitchcock EM, Warm JS, Matthews G, et al. Automation cueing modulates cerebral blood flow and vigilance in a simulated air traffic control task. Theor Issues Ergonom Sci 2003; 4:89–112.
12. Jewett ME, Dijk DJ, Kronauer RE, Dinges DF. Dose-response relationship between sleep duration and human psychomotor vigilance and subjective alertness. Sleep 1999; 22:171–9.
13. Kane RL, Kay GG. Computer applications in neuropsychological assessment. In: Horton A, Wedding D, Webster J, eds. The neuropsychology handbook: New edition. 2nd ed. New York: Springer; 1997:359–92.
14. Karcz-Kubicha M, Antoniou K, Terasmaa A, et al. Involvement of adenosine A1 and A2A receptors in the motor effects of caffeine after its acute and chronic administration. Neuropsychopharmacol 2003; 28:1281–91.
15. Kelly TL, Neri DF, Grill JT, et al. Nonentrained circadian rhythms of melatonin in submariners scheduled to an 18-hour day. J Biol Rhythms 1999; 14:190–6.
16. Krueger JM, Obal F Jr, Fang J, et al. The role of cytokines in physiological sleep regulation. Ann NY Acad Sci 2001; 933: 211–21.
17. Lindskog M, Svenningsson P, Pozzi L, et al. Involvement of DARPP-32 phosphorylation in the stimulant action of caffeine. Nature 2002; 418:774–7.
18. Mumaw RJ, Roth EM, Vicente KJ, Burns CM. There is more to monitoring a nuclear power plant than meets the eye. Hum Factors 2000; 42:36–55.
19. Neri DF, Oyung RL, Colletti LM, et al. Controlled breaks as a fatigue countermeasure on the flight deck. Aviat Space Environ Med 2002; 73:654–64.
20. Reifman J. Alternative methods for modeling fatigue and performance. Aviat Space Environ Med 2004; 75(3, Suppl.):A173–80.

21. Russo M, Thomas M, Thorne D, et al. Oculomotor impairment during chronic partial sleep deprivation. Clin Neurophysiol 2003; 114:723–36.
22. Schuff N, Amend DL, Knowlton R, et al. Age-related metabolite changes and volume loss in the hippocampus by magnetic resonance spectroscopy and imaging. Neurobiol Aging 1999; 20:279–85.
23. Tangney J. Quantifying human performance: A scientific method allows prediction of human operational performance. 2000; Retrieved from http://www.afrlhorizons.com/Briefs/Sept01/OSR0005.html. Verified 1/22/04.
24. Thomas JR, Ahlers ST, House JF, Schrot J. Repeated exposure to moderate cold impairs matching-to-sample performance. Aviat Space Environ Med 1989; 60:1063–7.
25. Tripp LD, Chelette T, Savul S, Widman RA. Female exposure to high G: Effects of simulated combat sorties on cerebral and arterial O_2 saturation. Aviat Space Environ Med 1998; 69:869–74.
26. Van Dongen HPA, Maislin G, Mullington JM, Dinges DF. The cumulative cost of additional wakefulness: Dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. Sleep 2003; 26:117–26.