Gluck, K. A., Gunzelmann, G., & Krusmark, M. A. (2009). Do the details matter? Comparing performance forecasts from two computational theories of fatigue. In A. Howes, D. Peebles, & R. Cooper (Eds.), *Proceedings of the Ninth International Conference on Cognitive Modeling*. Manchester, UK: University of Manchester.

Do the Details Matter?
Comparing Performance Forecasts from Two Computational Theories of Fatigue

Kevin A. Gluck (kevin.gluck@us.af.mil)
Glenn Gunzelmann (glenn.gunzelmann@us.af.mil)
Air Force Research Laboratory, 6030 S. Kent St.
Mesa, AZ, 85212 USA

Michael A. Krusmark (michael.krusmark@mesa.afmc.af.mil)
L-3 Communications, 6030 S. Kent St.
Mesa, AZ, 85212 USA

Keywords: Computational Cognitive Modeling; Fatigue; Alertness; Performance Forecasts; Comparison

A Tale of Two Theories

Activation and Micro-Lapses

We have been developing a computational theory of the effects of fatigue (especially sleep-related fluctuations in alertness) on the human cognitive system, implemented through mechanisms that impact existing components of the ACT-R architecture (Gunzelmann, Gluck, Kershner, Van Dongen, & Dinges, 2007; Gunzelmann, Gross, Gluck, & Dinges, 2009). These mechanisms include the suppression of *activation* in the declarative knowledge system, as well as brief breakdowns in the central production execution cycle, which we call *micro-lapses*.

Through an iterative series of mechanistic architectural modifications, model implementations, and goodness-of-fit evaluations in task contexts like the Psychomotor Vigilance Test (PVT – Dinges & Powell, 1985) and the Walter Reed Serial Addition/Subtraction Task (SAST – Thorne, Genser, Sing, & Hegge, 1985), the theory has evolved to a state in which we have some confidence in its appropriateness. In other words, we feel increasingly confident that the mechanisms we are using to replicate and explain relevant empirical results are both sufficient and necessary for that purpose (Estes, 2002). This gives us a measure of confidence that it is reasonable, perhaps even advisable, to use the theory to make novel performance predictions in task contexts beyond those used for originally developing and evaluating the theory. So far, we have promising results from fatigued performance predictions in both the context of dual-tasking (Gunzelmann, Byrne, Gluck, & Moore, 2009) and also in the context of simulated driving (Gunzelmann, Moore, Salvucci, & Gluck, submitted).

Cognitive Slowing

A popular alternative theory of fatigue is one commonly referred to as *cognitive slowing*. Though typically presented as a verbal-conceptual theory that describes an important category of empirical results from the sleep research community, cognitive slowing has inspired at least one prior computational implementation that explicitly moderated the processing of a simulated cognitive system by literally slowing it down (Jones, Laird, & Neville, 1998). To introduce fatigue effects into their model, Jones et al. modified Soar’s mechanisms to introduce artificial delays in processing, thereby having the effect of slowing overall system performance. Indeed, one of our very first conjectures regarding plausible mechanisms for implementing a theory of the effects of fatigue on cognitive processing involved a focus on “cognitive slowing” implemented as changes in the Default Action Time (DAT) of the production cycles in ACT-R, which controls the speed of central cognition in the architecture.

Does It Matter?

Despite what we consider to be convincing theoretical and empirical evidence that a cycle time-based account is less valid than our preferred “activation and micro-lapses” theory, we have been left to wonder whether the different theories would actually produce meaningfully different predictions in a more complex, dynamic, realistic context like aircraft maneuvering. This is more than just idle curiosity. It speaks to the core justification for pursuing basic computational cognitive modeling research – that the details matter – not only in the arena of theoretical constructs, but also in the arena of applied cognitive technologies.

Sleepy Pilot Performance Forecasts

We incorporated the fatigue mechanisms into a cognitive model that flies basic maneuvers with a Predator Synthetic Task Environment, in order to simulate the effects of extended sleep deprivation on pilot performance. Gluck, Ball, and Krusmark (2007) described the basic maneuvering task and cognitive model implementation in detail, and space considerations preclude repeating that material here. We will note, however, that for purposes of the pilot performance forecasts reported here we used Maneuver 7 (which requires simultaneous constant rate of change adjustments to airspeed, altitude, and heading over a 90-second trial) and we used the Control Focus and Performance variant of the pilot model, which is our most valid replication of expert-level performance on the basic maneuvering tasks.

With that model as a baseline, we implemented our set of mechanisms in to the model, and used parameter values derived from previous fatigue modeling efforts using the SAST, to arrive at principled values for the "Activation and Micro-Lapses" account. We also derived predicted DAT values for the "Cognitive Slowing" account using values estimated to account for dual-task performance. Though imperfect, the mechanisms and parameter values reflect an honest effort to faithfully implement and parameterize both accounts. To evaluate the alternatives, we ran the model 110 times at each of four levels of sleep deprivation: Baseline (no sleep deprivation), 1, 2, and 3 days of sleep deprivation. Results are presented in Figure 1.

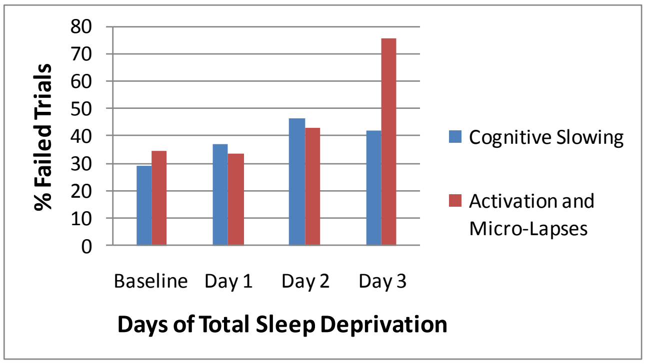

Figure 1: % failed basic maneuvering trials by fatigue theory, across four levels of sleep deprivation

The forecasts show nearly identical performance up to two days of sleep deprivation, followed by a dramatic difference in predicted performance level after three days without sleep. The obvious implication of this result is that it suggests that it *does* matter what the details are in your implementation of a theory of fatigue in the human cognitive system, at least in the extreme. However, this result also raises an assortment of more subtle issues associated with the challenges we face as we begin trying to make real, no kidding, a priori performance predictions in transfer contexts. Some of these questions include:

1. How sensitive are the predictions to variations in the model parameters?
2. How valid are the results?
3. Would we be comfortable using these results to inform policy decisions?

We hope to discuss and debate possible answers to these questions with attendees at ICCM 2009.

Discussion

The good news story is that we have reached a state in our research where we can make forecasts of this sort in complex, dynamic domains and have some confidence in the accuracy of those predictions. This is a desirable state for cognitive science in general, and for us in particular.

The bad news story is that we have no expectation of being able to directly evaluate the accuracy of the model predictions against empirical human data. It is logistically difficult and expensive to run the necessary sleep protocols with this task. It is an interesting conundrum that we are just beginning to face in computational cognitive science.

Acknowledgments

The views expressed in this paper 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 authors thank Kenny Jackson for his assistance with portions of this research. This 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 (AFOSR).

References

Estes, W. K. (2002). Traps in the route to models of memory and decision. *Psychonomic Bulletin & Review, 9*(1), 3-25.

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.

Gluck, K. A., Ball, J. T., & Krusmark, M. A. (2007). Cognitive control in a computational model of the Predator pilot. In W. Gray (Ed.), *Integrated models of cognitive systems* (pp. 13-28). New York, NY: Oxford University Press.

Gunzelmann, G., Byrne, M. D., Gluck, K. A., & Moore, L. R. (2009). Implications of computational cognitive modeling for understanding fatigue. *Human Factors.*

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 and J. G. Trafton (Eds.), *Proceedings of the Twenty-Ninth Annual Meeting of the Cognitive Science Society* (pp. 329-334). Austin, TX: Cognitive Science Society.

Gunzelmann, G., Gross, J. B., Gluck, K. A., & Dinges, D. F. (2009). Sleep deprivation and sustained attention performance: Integrating mathematical and cognitive modeling. *Cognitive Science*.

Gunzelmann, G., Moore, L. R., Salvucci, D. D., & Gluck, K. A. (submitted). Fluctuations in alertness and sustained attention: Predicting driver performance. Under consideration for presentation at the *9th International Conference on Cognitive Modeling*.

Jones, R., Laird, J., & Neville, K. (1998). Modeling pilot fatigue with a synthetic behavior model. *Proceedings of the Seventh Conference on Computer Generated Forces and Behavioral Representation*, Orlando, FL.

Thorne, D. R., Genser, S. G., Sing, G. C., & Hegge, F. W. (1985). The Walter Reed performance assessment battery. *Neurobehavioral Toxicology and Teratology, 7*, 415-418.