Examining Sources of Individual Variation in Sustained Attention

Glenn Gunzelmann (glenn.gunzelmann@mesa.afmc.af.mil)
Cognitive Models and Agents Branch, Air Force Research Laboratory
6030 South Kent St., Mesa, AZ  85212 USA

L. Richard Moore, Jr. (larry.moore@mesa.afmc.af.mil)
Lockheed Martin at Air Force Research Laboratory
6030 South Kent St., Mesa, AZ  85212 USA

Kevin A. Gluck (kevin.gluck@mesa.afmc.af.mil)
Cognitive Models and Agents Branch, Air Force Research Laboratory
6030 South Kent St., Mesa, AZ  85212 USA

Hans P. A. Van Dongen (hvd@wsu.edu)
Sleep and Performance Research Center, Washington State University Spokane
P.O. Box 1495, Spokane, WA  99210-1495, USA

David F. Dinges (dinges@mail.med.upenn.edu)
Division of Sleep and Chronobiology, University of Pennsylvania School of Medicine
1013 Blockley Hall; 423 Guardian Drive; Philadelphia, PA  19104-6021 USA

Abstract
Sustained attention and psychomotor reactions are foundational components of performance in many laboratory and applied tasks. In sleep research studies, individual differences in baseline attentional vigilance are compounded by individual differences in vulnerability to the negative consequences of fatigue due to sleep loss, producing large differences in reaction time profiles. In this paper, we present a theory and model to explain individual differences in reaction time performance in a sustained attention task, both at baseline and as overall alertness declines across 88 hrs without sleep. The model captures the performance of individual human participants, and illustrates how individual differences in processing speed and differences in susceptibility to fatigue from sleep loss may combine to produce unique performance profiles.

Keywords: Attention; Reaction Time; Individual Differences; Processing Speed; Computational Model.

Introduction
Attentional vigilance refers to the ability to maintain focused attention on a task and respond appropriately to repetitive stimuli. Vigilance is critical in monitoring tasks that are central in many transportation domains (e.g., train operators or long-haul truck drivers), and in many security-related tasks (e.g., baggage screeners and intelligence analysts). A substantial body of literature has accrued on breakdowns in attentional processes (e.g., Davies & Parasuraman, 1982; Van Dongen & Dinges, 2005), which can have serious consequences in applied settings (e.g., Caldwell, Caldwell, Brown, & Smith, 2004). Lapses in attention have been attributed to fatigue caused by sleep loss (e.g., Doran, Van Dongen, & Dinges, 2001; Dorrian, Rogers, & Dinges, 2005) and/or extended time on task (e.g., Davies & Parasuraman, 1982; Van Dongen & Belenky, 2008).

In the Psychomotor Vigilance Test, or PVT (Dinges & Powell, 1985; Dorrian et al., 2005), participants monitor a known location on a computer screen and press a response button each time a stimulus appears at that location, which happens at random intervals between 2 s and 10 s. Sustained attention is taxed in this task as a function of the length of each test session, which was fixed at 10 minutes for the experiment described below.

There are baseline differences among individuals in the speed with which they are able to respond to stimuli in reaction time tasks like the PVT (e.g., Humphreys & Revelle, 1984). Such individual differences in reaction time performance have been studied in the context of the relationship to general intelligence (e.g., Deary, Der, & Ford, 2001; Larson & Alderton, 1990), and explained in terms of processing speed, with slower processing being associated with both longer reaction times and lower overall intelligence.

In addition to differences in reaction times across individuals on the PVT and many other reaction time tasks, there are considerable differences in how reduced alertness resulting from fatigue impacts performance. Alertness in this context refers to overall cognitive performance capability, which varies as a function of time awake and circadian rhythms. Research on sleep deprivation has demonstrated substantial declines in performance on the PVT as a function of these factors (e.g., Doran et al., 2001; Dorrian et al., 2005; Van Dongen & Dinges, 2005). The extent of those declines varies significantly across
individuals and reflects a trait (Van Dongen, Baynard, Maislin, & Dinges, 2004).

We have explored computational mechanisms to explain individual differences in human performance on the PVT (Gunzelmann, Moore, Gluck, Van Dongen, & Dinges, 2008), as well as changes observed as alertness varies (e.g., Gunzelmann, Gross, Gluck, & Dinges, in press). Here we present an integrated account of PVT performance, which explains stable individual differences in performance through variations in processing speed, combined with distinct mechanisms to represent the deleterious impact of sleep deprivation. The resulting model provides a more comprehensive explanation of sustained attention performance and adds new insights regarding the nature of performance differences across individuals at baseline and over the course of an extended period without sleep.

**Model and Mechanisms**

Our computational model for the PVT was developed using the ACT-R cognitive architecture (Anderson et al., 2004). The PVT places emphasis on ACT-R’s perceptual and motor capabilities, which must encode the stimulus when it is presented and elicit a response efficiently to produce effective task performance. The coordination of these activities is accomplished by ACT-R’s central cognitive process, which is implemented as a serial production system that operates in a cyclical manner to represent goal-directed cognitive activity.

The foundation of the model consists of processes that (1) shift visual attention to the stimulus when it appears and (2) generate a response in the form of a virtual button press. These processes are represented as productions in ACT-R. The first process is sensitive to the appearance of the stimulus and generates a request for ACT-R’s visual system to shift attention to the item. The second process generates a response through a request to ACT-R’s motor system. Responses also can be generated in the absence of the stimulus, creating the possibility of false starts (see Gunzelmann et al., in press). Baseline differences and declines associated with fatigue are instantiated in the model through parameter manipulations that influence the duration and probability of successfully executing these processes. These mechanisms are described in the next subsections.

**Variability in Baseline Reaction Time**

Accounts of differences in reaction time implicate processing speed as the main factor. We represent this in the current model using a parameter in ACT-R that controls the duration of cognitive actions. Specifically, the parameter controls the time required for a single cognitive cycle within ACT-R’s central production system, which involves matching, selecting, and executing (firing) a single production. The default time for this process in ACT-R is 50 ms. In the model, noise is added to this parameter to produce variability in the timing of cognitive cycles. The noise is sampled from a uniform distribution ranging from 2/3 to 4/3 of the parameter value.

We manipulate the parameter controlling the duration of cognitive cycles to represent stable processing speed differences among participants. The parameter has two specific effects on the model’s performance. First, and most obviously, it has a direct impact on the mean time required to complete the task. By decreasing or increasing cognitive cycle time, the model becomes faster or slower in responding to the presentation of the stimulus on average. Second, because the width of the uniform distribution determining the variability in the timing of cognitive cycles is defined to be proportional to the cognitive cycle time, faster cognitive cycle times produce narrower distributions than longer cycle times. This predicts that individuals who are slower in performing the task will also be more variable in their reaction times.

**Performance Decrement with Decreased Alertness**

The mechanism responsible for individual differences in reacting to the onset of a stimulus under baseline conditions represents one aspect of the research presented here. The other aspect relates to individual differences in the ability to maintain performance on the task despite reductions in overall cognitive alertness stemming from extended periods of sleep deprivation. In our computational model, the impact of sleep deprivation on PVT performance is driven by mechanisms within the central production system.

The mechanisms allow for very brief gaps in cognitive processing, which we refer to as micro-lapses (Gunzelmann et al., in press). These micro-lapses reflect cognitive cycles in ACT-R where no cognitive actions are performed. As alertness declines, the likelihood of a micro-lapse increases, leading to delayed responses (lapses) and occasional failures to respond (non-responses). In ACT-R, the selection and execution of actions in central cognition is managed by the calculation of an expected utility for each production ($U_i$), which is influenced by an anticipated cost ($C_i$), a likelihood of success ($P_i$), and an overall level of “alertness” in the cognitive system ($G$). The equation for the expected utility of a production, $i$, is:

$$U_i = P_iG - C_i + \varepsilon$$

Note that noise ($\varepsilon$) is added to the utility computation, which allows for stochasticity in the selection and execution of cognitive actions. Micro-lapses occur in our model when none of the expected utilities for applicable productions exceed a threshold for action, referred to as the utility threshold ($T_u$). In this circumstance, no action is performed on that cognitive cycle, and it is followed by another cognitive cycle where utility values are evaluated once again to determine if an action will be executed. Noise in the utility computation, sampled from a distribution with a mean of 0 and a standard deviation of about 0.453 (a default value in ACT-R), is critical in creating a circumstance where a micro-lapse can be followed by an appropriate cognitive action, allowing for the possibility for delayed responses (i.e., lapses).
Declines in alertness are represented by decreasing $G$, which is a global parameter that impacts the utility value of all productions. $G$ is decremented further during cognitive inactivity (i.e., during micro-lapses) to represent dynamic declines in alertness over time. $T_u$ is also decremented as alertness declines; it reflects compensatory effort on the part of the individual to offset the negative consequences of fatigue (see Gunzelmann et al., in press, for details). The overall impact of decreases in $T_u$ is to make it more likely that some action will be performed on a given cognitive cycle. In this model, lower $T_u$ values are a main contributor to increased numbers of false starts seen with sleep deprivation (Doran et al., 2001).

To reduce degrees of freedom in the assessment of values for $G$ and $T_u$, the dynamics of their changes are constrained by predictions of alertness from a published biomathematical model (Jewett & Kronauer, 1999) representing the interplay of sleep homeostasis and circadian rhythms on alertness (see Mallis, Mejdal, Nguyen, & Dinges for an overview of this class of model). A linear mapping of the alertness predictions to $G$ and $T_u$ provides an effective means of constraining the changes in these parameters in the model (Gunzelmann et al., in press).

**Comparison to Individual Human Performance**

**Human Experiment Protocol**

To evaluate the ability of our model to capture a breadth of individual performance, including wide variations in alertness among individuals, we used data from 13 participants who completed a study involving 88 hrs of total sleep deprivation (Doran et al., 2001). Participants completed a 10-minute PVT session every two hrs throughout the sleep deprivation period as part of a battery of cognitive tasks. Responses were classified as false starts if made before or within 150 ms of the stimulus presentation. Alert responses were considered to be responses between 150 ms and 500 ms, while longer responses were categorized as lapses. In cases where no response was made within 30 s of the stimulus onset, the trial was halted and identified as a non-response while a beep alerted the participant for the next trial.

There were substantial inter-individual differences in performance overall, and the extended period of 88 hrs without sleep introduced wide temporal variations in alertness. In the current paper, we focus on declines in performance that occurred over progressive days without sleep while averaging out changes within days. Elsewhere we have used our approach to look at changes that occur across hrs within a day as a function of circadian rhythms (Gunzelmann et al., in press).

**Model Fitting and Evaluation**

For each individual, we explored the capacity of the model to capture average human performance for each day of the sleep deprivation protocol, including the baseline day that followed a full 8 hrs in bed and the subsequent first, second and third days of total sleep deprivation. The qualitative dynamics of the computational model were constrained by biomathematical model predictions of alertness, but we allowed magnitudes to vary on an individual basis. For every participant, we estimated intercepts and slopes to map the values of $G$ and $T_u$ to the biomathematical model predictions of alertness. Baseline cognitive cycle time was also estimated for every participant, but not varied as a function of predicted alertness because no such relationship was found ($p>.90$).

We based the evaluation of our model on a “standard two-stage” method. In the first stage of our analysis, we fitted the 5 parameters identified above for each participant. We then compared the model results to the human data by computing the proportion of responses classified as false starts, lapses, and non-responses, as well as proportions of responses falling into 10 ms bins across the alert response time range (150–500 ms). In the second stage, we computed summary statistics and based our conclusions on the behavior of the model across the whole sample. In this manner, we avoided overparametrization of the research problem—standard two-stage methods constitute a statistically appropriate and approximately valid approach to the study of individual differences (Feldman, 1988; Van Dongen, Maislin, & Dinges, 2004).

**Results**

Figure 1 illustrates the ability of the model to capture the range of human behavior, both at baseline and across an extended period of sleep deprivation. Increases in cognitive cycle time in the model produced shifts in the response distribution to the right, combined with a widening of the distribution attributable to noise. This prediction of the model is borne out in the human data. In fitting individual human performance data for the PVT, we found that the best-fitting values for cognitive cycle time ranged from 21ms to 70ms, which is largely in line with proposals made by Card, Moran, and Newell (1983) regarding individual variability in cognitive processing speed. Importantly, this single parameter did an excellent job of accounting for individual differences in human performance at baseline.

There was limited evidence for a systematic increase in cognitive cycle time across the sleep deprivation period. An increase in this parameter was supported by 7 of the 13 participants (none of which were statistically significant, $p>.08$), while 2 of the remaining 6 showed a significant trend in the opposite direction ($p<.05$). Overall, changes in cognitive cycle time did not lead to significant improvement in the model’s predictions ($p>.90$) These results offer further support for holding cognitive cycle time constant for each individual across time awake, and call into question cognitive slowing as the sole explanation for the impact of sleep deprivation on performance (see also Dinges & Kribbs, 1991).
In addition to cognitive cycle time, alertness ($G$) and the threshold for action ($T_u$) were fitted. There were substantial differences among participants in the baseline values of $G$ and $T_u$, and in the slopes of change for $G$ and $T_u$ as a function of the alertness changes across days of sleep deprivation. As a result, the values of $G$ and $T_u$ became more different among participants as sleep deprivation progressed.

For all 13 participants, performance was best fit when $G$ values declined across the sleep deprivation period in parallel with the biomathematical model predictions of alertness. Overall, the impact of changes to $G$ across the sleep deprivation period was significant ($p<.001$). A similar pattern was observed for $T_u$, where the data from 11 of the participants was best fit when $T_u$ declined as time awake increased. This effect was significant as well ($p<.001$). Table 1 presents the baseline value for each of the parameters, and the coefficient relating $G$ and $T_u$ to alertness, for the fits presented in Figure 1.

Table 1: Baseline values for all parameters and regression coefficients ($\beta$) to map $G$ and $T_u$ to the Jewett & Kronauer (1999) model for each participant, along with the correlation (R) and Root Mean Squared Deviation (D) of the model with the individual participant data in Figure 1.

<table>
<thead>
<tr>
<th>ID</th>
<th>$G$ Value ($\beta$)</th>
<th>$T_u$ Value ($\beta$)</th>
<th>Cycle (ms)</th>
<th>R</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.27 (1.32)</td>
<td>1.53 (.85)</td>
<td>38</td>
<td>.95</td>
<td>.013</td>
</tr>
<tr>
<td>B</td>
<td>1.51 (0.64)</td>
<td>1.70 (.23)</td>
<td>59</td>
<td>.97</td>
<td>.014</td>
</tr>
<tr>
<td>C</td>
<td>1.58 (1.69)</td>
<td>1.51 (1.26)</td>
<td>38</td>
<td>.92</td>
<td>.013</td>
</tr>
<tr>
<td>D</td>
<td>0.65 (3.12)</td>
<td>1.05 (2.50)</td>
<td>41</td>
<td>.90</td>
<td>.026</td>
</tr>
<tr>
<td>E</td>
<td>1.22 (0.75)</td>
<td>1.23 (.49)</td>
<td>70</td>
<td>.98</td>
<td>.010</td>
</tr>
<tr>
<td>F</td>
<td>1.68 (0.43)</td>
<td>1.68 (.21)</td>
<td>57</td>
<td>.95</td>
<td>.016</td>
</tr>
<tr>
<td>G</td>
<td>1.24 (3.07)</td>
<td>1.21(2.74)</td>
<td>70</td>
<td>.98</td>
<td>.010</td>
</tr>
<tr>
<td>H</td>
<td>1.63 (0.33)</td>
<td>1.68 (.04)</td>
<td>47</td>
<td>.94</td>
<td>.013</td>
</tr>
<tr>
<td>I</td>
<td>1.18 (2.23)</td>
<td>1.21 (2.06)</td>
<td>51</td>
<td>.92</td>
<td>.016</td>
</tr>
<tr>
<td>J</td>
<td>1.61 (0.56)</td>
<td>1.88 (.08)</td>
<td>24</td>
<td>.93</td>
<td>.014</td>
</tr>
<tr>
<td>K</td>
<td>2.18 (0.29)</td>
<td>1.96 (-.56)</td>
<td>21</td>
<td>.97</td>
<td>.012</td>
</tr>
<tr>
<td>L</td>
<td>1.58 (0.02)</td>
<td>1.71 (-.08)</td>
<td>37</td>
<td>.96</td>
<td>.011</td>
</tr>
<tr>
<td>M</td>
<td>1.33 (0.85)</td>
<td>1.54 (.41)</td>
<td>40</td>
<td>.97</td>
<td>.010</td>
</tr>
</tbody>
</table>

To evaluate the overall capacity of the model to capture human performance, aggregate statistics are presented in Table 2. Mean parameter values and standard deviations are shown to illustrate the variation required to capture behavioral differences observed across individuals. In addition, the means and the standard deviations of the correlation and root mean squared deviation (RMSD) values of the fits to the individual human data are presented. The relatively high average correlation and correspondingly low RMSD illustrate the model’s overall ability to capture individual-level performance well, while the low standard deviation of these statistics indicates that the model is generally effective for each of the individual participants modeled, as illustrated in Figure 1 and Table 1.
Table 2: Summary statistics of the model’s performance across individuals, including measures of parameter variation (ms) and fit to human data (correlation [r] and root mean squared deviation [RMSD]). RMSD is presented as a proportion of responses.

<table>
<thead>
<tr>
<th>Parameter/Statistic</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>G (Intercept)</td>
<td>1.438</td>
<td>0.358</td>
</tr>
<tr>
<td>G (Slope)</td>
<td>1.178</td>
<td>1.046</td>
</tr>
<tr>
<td>T_u (Intercept)</td>
<td>1.532</td>
<td>0.280</td>
</tr>
<tr>
<td>T_u (Slope)</td>
<td>0.787</td>
<td>1.045</td>
</tr>
<tr>
<td>Cycle Time</td>
<td>45.566</td>
<td>15.376</td>
</tr>
<tr>
<td>Correlation (r)</td>
<td>0.948</td>
<td>0.026</td>
</tr>
<tr>
<td>RMSD</td>
<td>0.014</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Conclusions

The results presented in this paper raise a number of issues with regard to understanding psychomotor vigilance and inter-individual differences in human reaction time, as well as variability in the impact of sleep loss. Three primary conclusions can be drawn. First, cognitive cycle time in ACT-R provides a useful way of understanding stable individual differences in baseline reaction time performance on the PVT. This aspect of the architecture impacts the rate of cognitive activity across contexts, providing a fairly direct instantiation of processing speed. The performance of our model supports the idea that individual differences in reaction time performance can be captured in a relatively direct manner in ACT-R by using the cognitive cycle time parameters, and provides a detailed, process-level account of the phenomena observed in human performance.

Second, our research shows that micro-lapses in cognitive processing can provide a parsimonious account of both delayed responses (i.e., lapses) and smaller shifts in the speed of alert reaction times. As such, micro-lapses may be the right computational model equivalent of the “wake state instability” phenomenon that has been proposed to underlie the shifts in the reaction time distribution which give rise to slower responses and lapses (Doran et al., 2001).

Third, changes in cognitive cycle time alone did not capture performance changes associated with sleep loss. This calls into question the construct of cognitive slowing as a catastrophic error by a specific individual based upon his or her cognitive capabilities and limitations, including a catastrophic error by a specific individual based upon his or her cognitive capabilities and limitations, including performance degradations associated with sleep deprivation. This has the potential to increase safety across myriad real-world domains.

Interestingly, vulnerability to the negative consequences of sleep deprivation on cognitive performance showed only a modest relationship to individual differences in baseline performance, as has been observed previously (Van Dongen et al., 2004). The correlation between cognitive cycle time and the magnitude of changes to both G and T_u across successive days without sleep was not significant (p > .16; r = .30 with G and r = .41 with T_u), suggesting that these two sources of individual differences represent relatively distinct influences on cognitive processing.

The free parameters in our modeling effort reflect claims about the underlying sources of individual differences in human performance on this task, and so they were expected to vary among individuals. Research has shown repeatedly that people’s performance varies extensively across a wide variety of tasks in virtually every domain of psychological study. Our goal is to use laboratory tasks to generate a comprehensive model of the performance of individuals, providing a capacity to predict individual performance on applied tasks where data are difficult or impossible to collect (e.g., Gunzelmann & Gluck, in press). We comment on this long-term focus more in the remainder of the paper, which discusses future directions.

Applications and Future Directions

As we develop a more robust and detailed account of human cognitive performance and the various moderators that impact behavior, it should become increasingly possible to make predictions about the performance of individuals in novel task contexts. For instance, the research presented here provides evidence for variability in processing speed across individuals on a sustained attention task. Because our mechanisms are specified within a cognitive architecture, there is potential for using performance data from simple tasks like the PVT to generate predictions of performance for individuals in other tasks as a function of this variability. The same is true with regard to the mechanisms associated with changes in alertness. The next step in this process is to use these mechanisms to make such predictions in another task where we have data from the same participants. This will provide evidence regarding the promise of the methodology, and also will illustrate the utility of using a unified theory of cognition as a means of building a cumulative account of the impact of fatigue on cognitive performance.

The longer-term opportunity we see in this research is not in modeling laboratory tasks like the PVT per se. Rather, it is to use individually tailored parameter values derived from simple laboratory tasks to make specific predictions in more complicated, naturalistic task environments. A major achievement would be to be able to predict the likelihood of a catastrophic error by a specific individual based upon his or her cognitive capabilities and limitations, including performance degradations associated with sleep deprivation. This has the potential to increase safety across myriad real-world domains.

A further direction of this research is to understand in greater detail the dynamics of human sustained attention. Progress on this front depends on understanding another major influence on alertness, namely time on task. It is well established that performance on attention-demanding tasks tends to decline as the task is performed for greater lengths of time, a phenomenon referred to as the vigilance decrement (e.g., Davies & Parasuraman, 1982; Van Dongen & Belenky, 2008). Providing a unified account of the
relationships among time awake, circadian rhythms, and time on task represents a major subgoal in being able to predict variability in human performance across time. The current line of research represents significant progress toward that goal.

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. 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). Human data collection was sponsored partly by AFOSR grants F49620-95-1-0388 and F49620-00-1-0266, and by NIH grant RR00040. HVD was supported by AFOSR grant FA9550-09-1-0136. DFD was supported by AFOSR grant FA9550-05-1-0293.

References


