Gamma Power as an Index of Sustained Attention in Simulated Vigilance Tasks

Taylor Curley (taylor.curley@cubic.com)
Cubic Defense
Beavercreek, OH 45324 USA

Lorraine Borghetti (lorraine.borghetti.1.ctr@us.af.mil)
Air Force Research Laboratory
WPAFB, OH 45433 USA

Megan B. Morris (megan.morris.3@us.af.mil)
Air Force Research Laboratory
WPAFB, OH 45433 USA

Abstract

Performance on the psychomotor vigilance test (PVT; Dinges and Powell, 1985)—a common index of sustained attention—is affected by the opposing forces of fatigue and sustained effort, where reaction times and error rates typically increase across trials and are sometimes offset by additional efforts deployed toward the end of the task (i.e., an "end-spurt"; c.f. Bergum and Klein, 1961). In ACT-R (Adaptive Control of Thought-Rational; Anderson et al., 2004), these influences on task performance have been modeled as latent variables that are inferred from performance (e.g., Jongman, 1998; Veksler and Gunzelmann, 2018) without connections to directly observable variables. We propose the use of frontal gamma (γ) spectral power as a direct measure of vigilant effort and demonstrate its efficacy in modeling performance on the PVT in both the aggregate and in individuals.

Keywords: ACT-R; EEG; fatigue; vigilance; microlapse

Introduction

A well-documented phenomenon in human performance research is the decline in performance during extended vigilance tasks due to cognitive and physical fatigue (c.f., Ackerman, 2011). The relative simplicity of common sustained attention tasks, such as the psychomotor vigilance test (PVT; Dinges and Powell, 1985), however, overshadows the complex and arcane connections between task outcomes and the neural mechanisms that give rise to these outcomes (Ishii et al., 2014; Kim et al., 2017). Despite this, changes in electroencephalographic (EEG) activity have been shown to provide a potentially reliable marker of mental fatigue (Tran et al., 2020).

One way to examine links between cognitive and neural mechanisms of sustained attention is by integrating data from behavioral and neural sources into a single model (Turner et al., 2017). In the ACT-R (Adaptive Control of Thought-Rational; Anderson et al., 2004) cognitive architecture, for example, researchers have begun to use event-related potentials (ERPs; Cassenti et al., 2011) and neural "blips" (Borst and Anderson, 2015) to link selection and duration of individual behaviors (productions) to EEG data. Despite extensive work on modeling the effects of time-on-task (Veksler and Gunzelmann, 2018) and sleep deprivation (Gunzelmann et al., 2009, 2015) on the PVT, ACT-R practitioners have yet to directly investigate the use of EEG in modeling fatigue-related decrements during vigilance tasks.

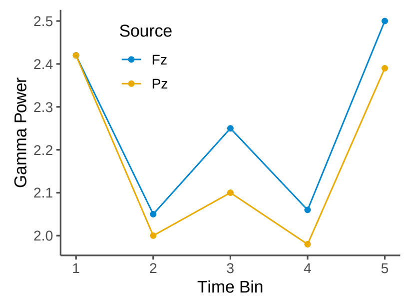

Figure 1: Observed relative gamma spectral power density across 2 minute time bins during the PVT. From Borghetti et al. (2021)[1].

We propose the use of estimated power in frontal gamma (γ) wave forms in models of vigilant attention. Specifically, we argue that γ power measured during the PVT is a reliable index of sustained attention that reflects fatigue (e.g., performance decreases across time) and compensation (e.g., end-spurts) and can be directly applied to ACT-R parameters. To this end, we first review relevant investigations of EEG and ACT-R as they relate to vigilance and then introduce a method for incorporating γ power into ACT-R models of the PVT.

EEG and Fatigue

Recently, Borghetti et al. (2021) reported a study examining electrophysiological measurements from 34 young adult participants ($M_{age} = 22.6$) over the course of a 10-min PVT in which participants were asked to respond immediately when a stimulus appears on the screen. Vigilance decrements during the PVT were exemplified by positive shifts in the distributions of reaction times, indicating increasingly slower responses, as well as increases in premature responses, i.e., false alarms (Doran et al., 2001). The results of the behavioral task also show a slight improvement in task performance in later trials, indicating an increase in effort, i.e. an "end-spurt" (e.g., Bergum and Klein, 1961).

The authors examined *spectral power density*, or an estimate of the power in a neural signal given a particular frequency, over the course of the 10-min task, focusing on theta (θ, 3-8 Hz), alpha (α, 9-14 Hz), beta (β, 15-30 Hz), and gamma (γ, 30-100 Hz) wave forms[1]. The top half of Figure 1 illustrates the main findings of the study: Significant trends indicating decreases in γ spectral power across time-on-task in both the frontal (Fz) and parietal (Pz) regions of the brain, with a significant end-spurt toward the end of the task (Morris et al., 2020). Borghetti et al. (2021) concluded that frontal γ indexes the dynamic between fatigue and sustained attention in the PVT. This is consistent with similar research indicating increases in γ activity across vigilance tasks (Kim et al., 2017) and positive associations between task performance and amplitudes of γ oscillations (Herrmann et al., 2010).

Fatigue and Compensatory Effort in ACT-R

The ACT-R cognitive architecture provides a rich environment for investigating effort and fatigue in goal-driven tasks, where influences on effort during the task are modeled as parameters affecting the selection and execution of procedural knowledge, i.e., "productions". During the course of the task, the model selects productions with the greatest estimated utilities (U), or a parameter indicating the strength and appropriateness of a given behavior at a given time. In prior versions of ACT-R, utilities were determined by the probability that a given goal will lead to success (P), the value of the current goal (G), and the cost of using that particular production to reach a goal (C). In the current version, production selection is a function of an initial utility value parameter (υ), noise on this value (σ^2), and a threshold parameter (τ), wherein the model selects the production with the highest above-threshold utility value to fire. Production utility values can either remain static or can update to reflect changes in the model's environment, such as production learning/reinforcement (e.g., Lovett and Anderson, 1996).

Previous studies have conceptualized vigilant effort as a direct influence on production utilities. Jongman (1998), for example, used parameterized "motivation" in a previous ACT-R architecture to directly influence G, where greater G values represent greater effort allocated toward achieving a goal and lead to better task outcomes, but lower G values result in firing inappropriate productions. Belavkin (2001) also used G to influence utility values, but conceptualized the parameter as reflecting a more general "arousal" state, where decreases in G result in fewer above-threshold productions, resulting in "giving-up" behavior. In contrast, Gunzelmann et al. (2009) simulated fatigue by imparting its effects on both utility values (through the G parameter) and τ as a function of "arousal" (A), which is derived from biomathematical estimates of arousal (c.f. Van Dongen, 2004). The decrease in utility and τ values represent the deleterious effects of fatigue and efforts enacted to compensate for fatigue, respectively. Gunzelmann et al. (2009) also incorporated "microlapses", or simulated lapses in attention. Microlapses occur when the utility module is unable to select a production, such as when all utility values are lower than τ. The occurrence of a microlapse results in a penalty to utility values and thus increases the probability of future microlapses. While the number of microlapses that occur during a simulated task is not controlled by the modeler, the penalty to utility values can be freely-estimated.

More recently, Veksler and Gunzelmann (2018) generalized decrements in arousal as stemming from the effects of time spent engaging in the task ("time-on-task") and simulated microlapses. Specifically, the authors estimate the utility of a production U by imposing a penalty on the initial production utility value (υ) as a function of both the number of microlapses (N_{ml}) and the time spent on the experiment (t):

$$U(t) = \upsilon\,[\lambda^{N_{ml}}(1+t)^{\rho}], \tag{1}$$

where υ is the initial utility value parameter, t is time spent on the task (scaled to minutes), λ scales the effect of microlapses on utility values, and ρ scales the effect of time-on-task. As fatigue increases and production values decrease, the probability of sampling an inappropriate production increases, leading to increases in false alarms.

In contrast, the production utility selection threshold is only affected by time-on-task:

$$UT(t) = \tau\,(1+t)^{\kappa}, \tag{2}$$

where τ is the initial utility theshold parameter and κ scales the effect of time-on-task on the threshold. Lower thresholds under conditions of fatigue allow the model to select productions whose υ values have decreased. This compensation is imperfect, however, as lowering the production selection threshold also allows the model to fire productions that are not appropriate for the context. In models of the PVT, this leads to increases in false starts and misses.

Candidates for Integration

We now review mechanisms for 1) translating γ spectral power to units appropriate for use in ACT-R simulations and 2) applying transformed γ estimates to the ACT-R cognitive architecture.

Scaling Spectral Power Estimates. Similar to previous research (e.g., Belavkin, 2001; Gunzelmann et al., 2009; Jongman, 1998; Veksler and Gunzelmann, 2018), we conceptualize sustained attention as a parameter ζ that is typically bounded between zero and one. In the proposed model, however, ζ can occasionally exceed its upper bound, meaning that parameterized effort cannot go below zero (meaning "absolute" fatigue), but can surpass unity (meaning "extra" effort). Thus, ζ can capture decrements due to fatigue as well as compensatory efforts that offset fatigue, such as the end-spurt effect (Morris et al., 2020).

[1] These results are based on a correction to the gamma spectral power analyses. In the original version of the paper, gamma estimates decreased sharply between 0 and 2 minutes and declined slightly across minutes 2 and 10. The corrected analyses indicate that gamma power increases between time bins 4 (6 - 8 m) and 5 (8 - 10 m), as shown in Figure 1.

Param.	Description	Bounds	Value	ζ Model?
υ	Initial production utility value	$[\,0.0, \text{Inf}\,]$	Free	Yes
τ	Initial production utility theshold	$[\,0.0, \text{Inf}\,]$	Free	Yes
ρ	Production utility time-on-task penalty	$[-1.0, 0.0\,]$	Free	No
κ	Utility threshold time-on-task penalty	$[-1.0, 0.0\,]$	Free	No
λ	Microlapse penalty	$[\,0.0, 1.0\,]$	Free	Yes
ϕ	Conflict resolution time	*N/A*	0.05	Yes

Table 1: Descriptions of fatigue-related parameters in the ACT-R model of the PVT. The "Value" column indicates if a value is freely-estimated, and if not, what the value is fixed to. The "ζ Model?" column indicates if the parameter is included in the model that uses γ power as a performance moderator. All 6 parameters are included in the full ("Fatigue") model (c.f. Veksler and Gunzelmann, 2018).

One way to normalize fatigue moderator values is by adjusting the values to the smallest value and the range of the values. This normalization method has been used to scale biomathematical estimates of arousal in previous investigations of the PVT (Gunzelmann et al., 2009), where estimates start with high values and monotonically decrease as a function of time. An interesting aspect of this method that is reflected in the fatigue moderators proposed by Gunzelmann and colleagues (Gunzelmann et al., 2009; Veksler and Gunzelmann, 2018) is that the normalized values start at 1 (the highest possible value) and decrease with time-on-task, implying that performance cannot meet or exceed that from $t = 1$. Therefore, we opted to normalize γ_i to the first observation in order to simulate end-spurt effects.

Given a set of observed spectral power estimates (total or relative) $\Gamma_i = \{\gamma_{i,1}, \ldots, \gamma_{i,t}\}$, for participant i at time t, as well as the range of these values, $\gamma_{r_i} = \text{range}\{\gamma_{i,1}, \ldots, \gamma_{i,t}\}$, we can calculate effort as:

$$\zeta_{i,t} = 1 + \left(\frac{\gamma_{i,t} - \gamma_{i,1}}{\gamma_{r_i}} \right). \tag{3}$$

Here, $\zeta_{i,1} = 1$ and all subsequent values are interpreted as diminished effort due to time-on-task ($\zeta_{i,t} \leq \zeta_{i,1}$) or additional (i.e., compensatory) effort compared to baseline ($\zeta_{i,t} \geq \zeta_{i,1}$), allowing the model to account for end-spurt effects.

Applying Fatigue Decrements. The theoretic interpretation of γ with respect to vigilance is intentionally vague (i.e., an index of sustained attention) and does not allow for a straightforward implementation of the ζ parameter in the ACT-R architecture. In these simulations, we integrate parameterized effort in a linear function with the initial production utility parameter υ (similar to Eq 1) with brief lapses in attention. Therefore, the modulated utility value at a given time, $U(t)$, can be calculated as a function of υ, ζ, and the number of simulated microlapses (N_{ml}):

$$U(t) = \upsilon \cdot [\lambda^{N_{ml}} \cdot \zeta_{i,t}]. \tag{4}$$

The Current Study

The estimated penalties to utility values and thresholds in ACT-R are imperfect. First, they are "smoothed" approximations of behavior and are unlikely to directly capture stochastic, asynchronous intraindividual variability across time, leading to error inflation when fitting fatigue parameters to individual participants. Second, these mechanisms are indirect inferences resulting from observations of behavioral data and have yet to be empirically linked to outside indicators.

The current project addresses these issues by examining the extent to which neural indices of vigilance correspond to the deleterious effects of fatigue in the PVT. Specifically, spectral power density in γ waveforms is expected to accurately capture fatigue and effort in ACT-R models of task performance. We expect to find that models using the observed power density estimates (Equations 3 and 4) in place of fatigue functions (Equations 1 and 2) will fit the observed data as well as, if not better than, models with these functions in both the aggregate and at the level of the individual.

Methods

Thirty-four adult volunteers (M_{age} = 22.60; SD_{age} = 4.08) recruited through the University of Dayton Research Institute (UDRI) participated in a single 2-h study session consisting of three experiment tasks with simultaneous EEG recording. The study was approved by institutional review boards at both UDRI and the Air Force Research Laboratory (AFRL), and all individuals were compensated for their participation in the study.

We provide a quick overview of the behavioral and electrophysiology methods below; further details can be found in Borghetti et al. (2021).

Behavioral

Participants were asked to participate in a 10-m PVT task as a part of the 2-h study session. During the PVT, participants were asked to monitor a computer screen with a black background and to press "j" on a standard computer keyboard as quickly as possible to a target stimulus, i.e., white numbers in the middle of the screen displaying the time (in ms) since target onset. The time in between the previous response and the onset of a new stimulus, the interstimulus interval (ISI), was randomly selected from an interval between 2 and 10 s. ISIs were exact integers and selected from a uniform distribution.

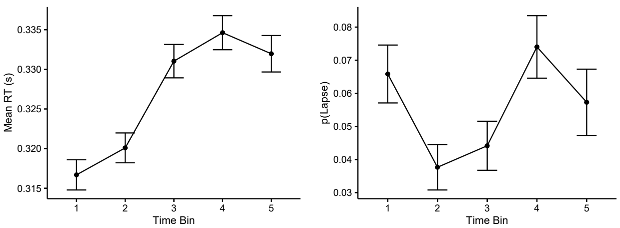

Figure 2: Performance data by time bins for average RTs for valid trials (left) and average proportion of lapses (right). Error bars represent the standard error of the mean.

EEG

Briefly, participants were fitted with an EEG cap with 64 electrodes, with 2 flat, unlinked electrodes applied to the mastoids. These data were processed using custom MATLAB scripts along with the EEGLAB toolbox (Delorme and Makeig, 2004). After applying a 1 Hz high-pass filter and removing artifacts, these data were epoched into segments of $\pm$1500 ms with respect to stimulus onset and divided into five, 2-m time bins. For the gamma spectral analysis, we assayed power in the 70-100 Hz frequency band for frontal (Fz) and parietal (Pz) cortical regions.

Computational

The computational model was programmed using a Julia language (Bezanson et al., 2017) implementation of the ACT-R cognitive architecture (Anderson et al., 2004). In ACT-R, the PVT has been modeled as a time-inhomogenous semi-Markov process consisting of three phases (Gunzelmann et al., 2009; Veksler and Gunzelmann, 2018): *Wait*, *Attend*, and *Respond*. The *Wait* production occurs prior to stimulus onset in anticipation of the next trial, while the *Attend* and *Respond* productions occur after a critical stimulus has been visually processed and after the decision has been made to engage in a response, respectively. These productions typically occur in the *Wait-Attend-Respond* sequence, but the order can be disrupted if an inappropriate production is selected on the basis of low utility values (U). This can lead to *false starts*, where the *Respond* production is selected in the absence of a valid stimulus (i.e., RTs $<$ 150 ms), and *lapses*, where the model fails to select the *Attend* or *Respond* productions in the presence of a valid stimulus (i.e., RTs $>$ 500 ms). Additionally, response latency is penalized whenever there are no productions that exceed the production utility threshold (UT) by adding 50 ms for each occurrence (*microlapse*; c.f. Gunzelmann et al., 2009).

Importantly, the ACT-R model of the PVT simulates fatigue by applying a penalty to a) only initial utility values (Belavkin, 2001; Jongman, 1998) or b) both initial utility values and utility thresholds (Gunzelmann et al., 2009, 2015; Veksler and Gunzelmann, 2018). Here, we only penalize utility values derived from Equations 3 and 4 based on re-scaled gamma power estimates. Table 1 provides descriptions of the parameters, the ranges of possible values, and the models that they are used in.

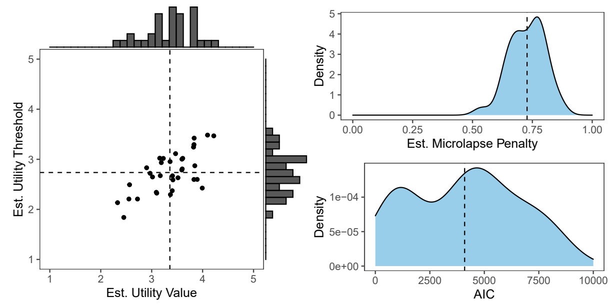

Figure 3: Best-fitting estimates for υ and τ (left), λ (top right), and associated AIC values (bottom right).

Results

Behavioral

We performed statistical analyses on responses categorized into 3 types: False starts (RTs $<$ 150 ms), lapses (RTs $>$ 500 ms), and valid responses (150 ms $\leq$ RTS $\leq$ 500 ms). For computational ease, we binned the data into five, 2-m bins and applied an inverse transformation to the RTs, i.e., $1/(RT * 1000)$ (Ratcliff, 1993).

A repeated measures ANOVA on the aggregated inverted RT values with a Greenhouse-Geisser correction on the degrees of freedom (W = 0.53, p = 0.02) indicates that the effect of time bin is significant, F(2.89,98.41) = 11.54, $p <$ 0.05, where average RTs increase between the first and fourth time bins (i.e., minutes 0 - 8), but decrease slightly in the fifth time bin (i.e., minutes 8 - 10; c.f. Figure 2). A similar one-way logistic GLM on lapses indicates that the log-odds of this type of response change across time bins, F(4,3652) = 3.48, $p <$ 0.05, where lapse rates decrease between bins 1 and 2, increase between bins 2 and 4, and then decrease again between bins 4 and 5 (c.f. Figure 2). A one-way logistic GLM indicates that the probability of a false start on any given trial is not different across time bins, F(4,3651) $<$ 0.1.

Spectral Power

For frontal γ (Figure 1), a Friedman test on total power estimates across time bins is significant, $\chi^2(4)$ = 11.3, p = 0.02. Follow-up paired comparisons indicate that estimates increase significantly between bins 2 and 3, $p <$ 0.05, decrease significantly between bins 3 and 4, $p <$ 0.05, and increase with marginal significance between bins 4 and 5, p = 0.07, although only the significance of the first comparison survives after Bonferroni corrections to the degrees of freedom.

Computational

We estimated the parameters for two different models—one using the fatigue moderators described by Veksler and Gunzelmann (2018) and another using gamma power estimates—using the data from individual participants and ag-

Estimate	Model	Parameters υ	τ	ρ	κ	λ	Fit Indices -2LL	AIC	BIC
Aggregate	Fatigue	4.01	2.90	-0.28	-0.20	0.98	5829.99	5839.99	5877.58
	Gamma	3.15	2.07	-	-	0.74	3920.76	3926.76	3949.25
Individual	Fatigue	5.78	0.32	-0.41	-0.17	0.81	6591.08	6601.08	6604.58
		(*0.51*)	(*0.08*)	(*0.05*)	(*0.03*)	(*0.03*)	(*227.69*)	(*227.69*)	(*227.69*)
	Gamma	3.36	2.74	-	-	0.73	4090.21	4096.21	4098.01
		(*0.08*)	(*0.07*)	-	-	(*0.01*)	(*424.64*)	(*424.64*)	(*424.64*)

Table 2: Best-fitting parameters for aggregated data (top) and summary statistics of the best-fitting parameters for individuals (bottom). For individuals, we report the means and standard errors of the mean (in parentheses) of these estimates. "Fatigue" refers to models using the decrement parameters described by Veksler and Gunzelmann (2018) while "Gamma" refers to the proposed model.

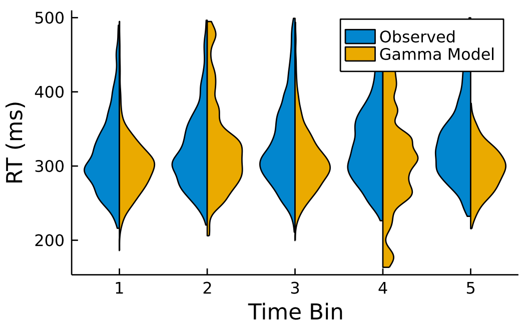

Figure 4: Reaction time distributions for valid responses across time bins for observed RTs (blue) and simulated RTs generated using the Gamma model (yellow).

gregated across all participants. Model fit was calculated using the summed log-likelihoods of the simulated RT data to log-normal distributions based on the observed RTs. We used a simplex search algorithm via `Optim.jl` (Mogensen and Riseth, 2018) to find the parameter values that maximized the likelihood of the two PVT models given the observed data. We repeated the optimization procedure 15 times for each set of data, using new starting values on each iteration to avoid local minima. Table 2 details the best-fitting parameters by data source ("Aggregate" vs. "Individual") and by the type of model ("Fatigue" vs. "Gamma").

Overall, the model using gamma spectral power density as a direct influence on utility values provides a better fit to the observed data than the model using established computational fatigue moderators. For the aggregated data, the difference in fit statistics suggest that there is decisive evidence (Kass and Raftery, 1995) in favor of the Gamma model, $\log B_{10}$ = 1928.33. Similarly, the difference in average fit values across all participants for the two models also suggests that there is decisive evidence in favor of the Gamma model, $\log B_{10}$ = 2506.57. Across individuals, the Gamma model is favored over the Fatigue model for all but 5 of the 34 participants in the study.

Discussion

In this paper, we introduced an ACT-R model of vigilant attention that directly integrates frontal γ spectral power density estimates into the parameters of the model that influence task performance. We compared the ability of the new model to fit observed RT data to that of a similar model of PVT performance and found that the proposed model provides a better fit to both aggregated and individual data than previous models of fatigue. These results suggest that frontal γ power estimates can be used as a measure of sustained attention and effort in models of vigilance.

The proposed model represents an initial step in developing models of fatigue and vigilance that incorporate directly-observable neural data. In this model, changes in observed neural data simply constrain the parameters of the behavioral model, i.e., a "direct-input approach" (Turner et al., 2017), implying a unidirectional influence. Future models, however, will need to simultaneously account for both neural and behavioral data and account for the bidirectional relationship between the two. Similarly, the use of frontal γ power in our model represents only one potential application of EEG data in cognitive models; our future research will use similar models to explore how other neural indices, such as beta (β) and alpha (α) frequency bands, can be used as observable estimates of fatigue and arousal in computational models of vigilance.

Acknowledgments

The opinions expressed herein are solely those of the authors and do not necessarily represent the opinions of the United States Government, the U.S. Department of Defense, the U.S. Air Force, or any of their subsidiaries, or employees. Distribution A. Approved for public release. Case number AFRL-2022-1771. The authors thank Bella Veksler and Chris Fisher for their help in model development and comments on the paper.

References

Ackerman, P. L. (2011). *Cognitive fatigue: Multidisciplinary perspectives on current research and future applications.* American Psychological Association.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., and Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, 111(4):1036–1060.

Belavkin, R. V. (2001). Modelling the inverted-U effect in ACT-R. In *Proceedings of the 2001 Fourth International Conference on Cognitive Modeling*, pages 275–76.

Bergum, B. O. and Klein, I. C. (1961). A survey and analysis of vigilance research. Technical report, No HUMRRO-RR-8. George Washington University, Alexandria, VA. Human Resources Research Office.

Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM Review*, 59(1):65–98.

Borghetti, L., Morris, M., Rhodes, L. J., Haubert, A., and Veksler, B. (2021). Gamma oscillations index sustained attention in a brief vigilance task. International Annual Meeting of the Human Factors and Ergonomics Society.

Borst, J. P. and Anderson, J. R. (2015). The discovery of processing stages: Analyzing EEG data with hidden semi-Markov models. *NeuroImage*, 108:60–73.

Cassenti, D. N., Kerick, S. E., and McDowell, K. (2011). Observing and modeling cognitive events through event-related potentials and ACT-R. *Cognitive Systems Research*, 12(1):56–65.

Delorme, A. and Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21.

Dinges, D. F. and 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.

Doran, S. M., Van Dongen, H. P., and Dinges, D. F. (2001). Sustained attention performance during sleep deprivation: evidence of state instability. *Archives of Italian Biology: Neuroscience*, 139(3):253–267.

Gunzelmann, G., Gross, J. B., Gluck, K. A., and Dinges, D. F. (2009). Sleep deprivation and sustained attention performance: Integrating mathematical and cognitive modeling. *Cognitive Science*, 33(5):880–910.

Gunzelmann, G., Veksler, B. Z., Walsh, M. M., and Gluck, K. A. (2015). Understanding and predicting the cognitive effects of sleep loss through simulation. *Translational Issues in Psychological Science*, 1(1):106–115.

Herrmann, C. S., Fründ, I., and Lenz, D. (2010). Human gamma-band activity: a review on cognitive and behavioral correlates and network models. *Neuroscience & Biobehavioral Reviews*, 34(7):981–992.

Ishii, A., Tanaka, M., and Watanabe, Y. (2014). Neural mechanisms of mental fatigue. *Reviews in the Neurosciences*, 25(4):469–479.

Jongman, G. (1998). How to fatigue ACT-R. In *Proceedings of the second European conference on cognitive modelling*, pages 52–57. Nottingham University Press.

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795.

Kim, J.-H., Kim, D.-W., and Im, C.-H. (2017). Brain areas responsible for vigilance: an eeg source imaging study. *Brain Topography*, 30(3):343–351.

Lovett, M. C. and Anderson, J. R. (1996). History of success and current context in problem solving: Combined influences on operator selection. *Cognitive Psychology*, 31(2):168–217.

Mogensen, P. K. and Riseth, A. N. (2018). Optim: A mathematical optimization package for Julia. *Journal of Open Source Software*, 3(24):615.

Morris, M. B., Haubert, A. R., and Gunzelmann, G. (2020). Beyond the vigilance end-spurt with event-related potentials. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 64, pages 1258–1262. SAGE Publications Sage CA: Los Angeles, CA.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, 114(3):510.

Tran, Y., Craig, A., Craig, R., Chai, R., and Nguyen, H. (2020). The influence of mental fatigue on brain activity: Evidence from a systematic review with meta-analyses. *Psychophysiology*, 57(5):e13554.

Turner, B. M., Forstmann, B. U., Love, B. C., Palmeri, T. J., and Van Maanen, L. (2017). Approaches to analysis in model-based cognitive neuroscience. *Journal of Mathematical Psychology*, 76:65–79.

Van Dongen, H. (2004). Comparison of mathematical model predictions to experimental data of fatigue and performance. *Aviation, Space, and Environmental Medicine*, 75(3):A15–A36.

Veksler, B. Z. and Gunzelmann, G. (2018). Functional equivalence of sleep loss and time on task effects in sustained attention. *Cognitive Science*, 42(2):600–632.