The Influence of Fatigue on Usage of Model-Based vs Model-Free Reinforcement Learning Strategies

Colton Loftus
Department of Computer Science, Princeton University

Maya Rozenshteyn
Department of Computer Science, Princeton University

Abstract

Cognitive models often assume that humans optimize tasks in a generally rational manner. Analyzing human task-related behavior under a state of fatigue, however, may challenge these assumptions of rationality. Physical and mental fatigue may alter not only task performance but also fundamental approaches to task-solving. This paper explores behavior during the two-stage Markov decision task completed under conditions of differing fatigue. Analysis of fatigue-modulated behavior in this task illustrates statistically significant differences in people's use of model-based versus model-free reinforcement learning strategies depending on fatigue level. Specifically, decreased fatigue results in decreased usage of both model-based and model-free strategies.

Keywords: two-stage decision task, reinforcement learning, model-based, model-free, fatigue

Introduction

Fatigue, a phenomenon that most people experience regularly, often has significant impacts on cognitive functioning. When fatigued, humans consistently perform significantly worse on a variety of different tasks than when alert and attentive. Beyond performance impacts, fatigue (given its effects on memory storage and retrieval, amongst other implications) may fundamentally alter the way humans approach and cognitively optimize various tasks. Many cognitive modeling-based studies do not factor the notion and impacts of fatigue into their modeling paradigms. Such studies often create a single model for behavior on a given task, neglecting to account for the likely many different cognitive states of the participants whose behavioral data the model is fitted to. Even models that account for "resource rationality" (the notion that humans have limited cognitive resources) still often aim to find the optimal biologically feasible mind, whereas the fatigued mind may be sub-optimal (even when accounting for general biological limitations). This paper aims to address this gap in existing modeling paradigms by assessing the impact of fatigue on people's use of model-based versus model-free reinforcement learning when evaluating candidate actions during a decision task.

Though the impact of fatigue on task performance is relatively intuitive and readily quantifiable, fatigue's influence on task approach and strategy is much more difficult to empirically measure. The neural correlates of relatively complex reinforcement learning strategies are unclear and explicit questioning of experiment participants on reinforcement learning strategy usage (i.e., asking participants questions along the lines of "did you use a model-based approach while performing this task") is unlikely to prove useful due to the technical expertise required to understand such strategies. As a result, this paper aimed to find a task well-suited to elucidating fatigue's influence on task approach. Such a task needed to be able to be taken at multiple different times (i.e., so that participants could complete the task at differing fatigue levels), produce quantifiable markers of task participants' behavior, yet also have enough probabilistic elements such that there would not be an immediately clear optimal solution (i.e., include a learning component).

To solve this issue, this study used a standardized two-stage Markov decision task (also known as the "two-step task") first proposed by Daw, Niv, and Dayan in 2005. The two-step task is an established experimental paradigm designed to isolate whether those who complete the task are using a model-based reinforcement learning strategy, a model-free reinforcement learning strategy, or some combination of the two. A key variation between the original task paradigm and the task paradigm used in this study is that the participants involved in this study were asked to complete the two-step task at two separate times: once at midday and once at night. Running the task on participants with differing fatigue levels (due to the different times of day at which participants completed the task) isolates the influence of fatigue on reinforcement learning strategy usage.

By running the two-step task with a group of individuals during both daytime and nighttime, we were able to use mixed-effects logistic regression to determine the impact of fatigue on the use of model-based versus model-free reinforcement learning strategies.

Background

There has been extensive research regarding the effects of fatigue on cognitive function, as well as extensive research modeling human cognitive behavior as reinforcement learning. This study seizes an opportunity to unite these two research domains and specifically focuses on the effect of fatigue on the use of model-based versus model-free reinforcement learning approaches when determining the optimal action to perform during a decision task.

Fatigue's Impact on Cognition

The influence of tiredness (physical and cognitive fatigue) on behavior has been studied extensively in the realms of athletics (Almonroeder, Tighe, Miller, & Lanning, 2018), economics (Mullette-Gillman, Leong, & Kurnianingsih, 2015), and job performance (Persson, Barrafrem, Meunier, & Tinghög, 2019), given its far-reaching impacts. All of the aforementioned studies have found significant decreases in performance as a result of fatigue and/or sleep deprivation. The cognitive correlates of fatigue, however, have proven difficult to quantify precisely. Some researchers postulate that fatigue results in an inflated cost of cognitive control, exemplified by more impulsive decisions (and have modeled fatigue in this manner, with limited results) (Wiehler, Branzoli, Adanyeguh, Mochel, & Pessiglione, 2022). In other terms, fatigued decision-makers are seen as optimizing a simpler function than energized decision-makers, given the increased cost of cognitive exertion when fatigued.

Gaps in Pre-existing Cognitive Science Models

Oftentimes, cognitive models assume that individuals are generally behaving rationally. In Bayesian modeling, researchers often assume that participants are capable of performing Bayesian inference with reasonable priors given the true state of the world (Khalvati et al., 2019). While humans don't have unlimited cognitive resources, cognitive models often assume humans are generally rationally optimizing tasks. Thus, it is to be expected that people participating in a task will eventually arrive at a more optimal solution as their priors become more well-informed. Even "resource-rational" cognitive models, which take into account human cognitive limitations and consider deviations from rationality, often assume a "bounded optimal mind" and may have a difficult time precisely incorporating the cognitive computational cost of atypical cognitive states and capacities (such as is often the case when a mind is fatigued) (Griffiths, Lieder, & Goodman, 2015).

Given the aforementioned literature on the behavioral effects of fatigue (which often leads to less-than-rational behavior), we want to examine how these general assumptions of rationality change when individuals are experiencing fatigue and are no longer at their optimal cognitive capacity. Specifically, we aim to examine the effects of fatigue on people's use of model-based versus model-free reinforcement learning during a decision-making task. Does people's behavior trend toward an optimal learning paradigm, given the resource-constrained nature of a fatigued mental state?

Model-based reinforcement learning has been established to be more cognitively demanding than model-free reinforcement learning (as model-based reinforcement learning involves re-computing a value function) and fatigue has been established as a major factor in the perception of cognitive demand (Dezfouli & Balleine, 2013). As such, analyzing the nature of people's reinforcement learning paradigm under differing levels of fatigue is particularly appropriate for isolating the effects of fatigue on cognition.

Background Summary

There is significant literature illustrating the impact of fatigue on cognitive functioning and the deviations from rationality fatigue may cause. Despite this, many cognitive modeling paradigms do not factor in notions of fatigue, operating under the assumption that agents are generally seeking to maximize reward and are capable of doing so rationally. In addition, while the impact of fatigue on performance has been well studied, the impact of fatigue on cognitive mechanisms is less clear. This paper seeks to remedy these notions by examining the impact of fatigue on the use of model-based versus model-free reinforcement learning strategies during a decision task. Does fatigue affect the rate at which people use either model-based or model-free learning strategies? By sampling both fatigued individuals and those who are well-energized, can we uncover a relationship between learning strategy and energy level?

Approach

Previous research has elucidated how a sequential choice task can be used to deduce if an agent is using a model-based or model-free reinforcement learning approach to make decisions (N. D. Daw, Gershman, Seymour, Dayan, & Dolan, 2011). The two-step decision task is well-documented in cognitive science literature (N. Daw, Niv, & Dayan, 2005), so this paper will only explain it briefly in Figure 1. The probabilistic nature of the two-step decision task means that different strategies for reinforcement learning predict different patterns by which reward obtained in the second stage should impact first-stage choices on subsequent trials (Ibid). For instance, if the task gives an unlikely transition from the first to the second stage and then gives a positive reward, a model-free learner would increase their probability of selecting the same first-stage choice, regardless of first-to-second stage transition probabilities. However, a model-based learner would recognize this irregularity and that "any increase in the value of the rewarded second-stage option will more greatly increase the expected value of the first-stage option that is more likely to lead there" (Ibid).

As the fatigue of two-step task participants can be correlated with their use of different reinforcement learning strategies, the sequential choice paradigm is particularly appropriate for deducing the implications of fatigue on decision-making. We build upon the sequential choice task framework, factoring in external circumstances (i.e., fatigue) into our analysis of an agent's ability to identify patterns and maximize reward.

Methods

Participants

This study included $n = 16$ participants (mean age ≈ 29 years). Participants were recruited via email (contacts of the

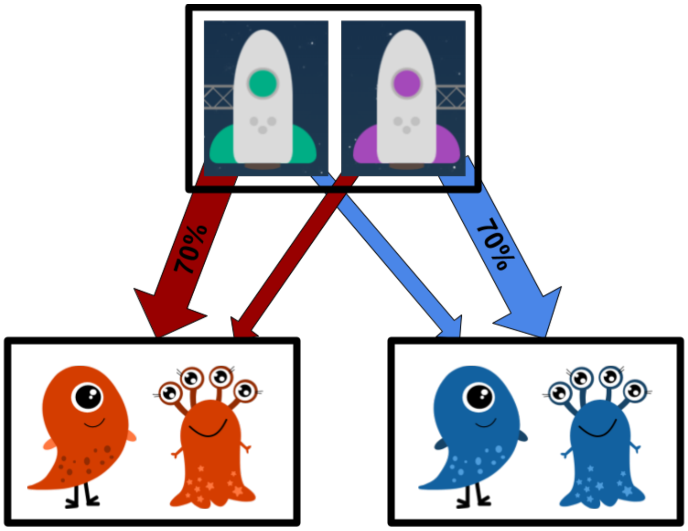

Figure 1: Two-step task state transition structure. Each first-stage choice has a 70% probability of transitioning to one pair of second-stage states and a 30% probability of transitioning to the other. Each second-stage choice is associated with a probability of obtaining a binary reward. Figure adapted from Daw, Gershman, Seymour, Dayan, and Dolan, 2011.

researchers). Each participant performed two trials, producing $n = 32$ total datasets for analysis. Each participant was assigned to one of two experimental conditions. The "day first" condition involved completing the first experimental trial between 11:00 am and 2:00 pm and the second experimental trial between 8:00 pm and 11:00 pm (on the same day). The "night first" condition involved completing the first experimental trial between 8:00 pm and 11:00 pm and the second experimental trial between 11:00 am and 2:00 pm (on the following day). There were an equal number of participants assigned to each condition ($n = 8$).

Materials

Our task involved exclusively web-based materials. Specifically, participants had to fill out a "tiredness assessment" hosted on Google Forms[1] and a version of the two-step task hosted on GitHub Pages.[2] Participants did not require any special materials or other software. The task was standardized for all computers given its web application format.

The two-step modeling task used in this study is adapted from the jspsych implementation of the two-step task from the Princeton University NivLab (authors: Sam Zorowitz, Gili Karni, Branson Byers).[3] We modified the original codebase to include features relevant to our experiment, such as exporting the time of day, altering the trial length, and revising task instructions.

Procedure

To complete our experiment, participants first had to fill out the aforementioned "tiredness assessment." The assessment used the "Stanford Sleepiness Scale" to gauge participants' levels of fatigue. The Stanford Sleepiness Scale has been established as an effective measure of subjective tiredness (Shahid, Wilkinson, Marcu, & Shapiro, 2011)

Immediately following the completion of the tiredness assessment, participants had to complete the aforementioned online two-step task. The version of the two-step task used here is set in outer space. After running through a set of task instructions, participants run through a series of $n = 150$ task runs (split into two sections of $n = 75$ task runs with a short break in between). For each task run, participants must choose either a green or purple rocket ship (the first decision level in the task). As depicted in Figure 1, each ship is probabilistically favored to go to a planet with aliens of a certain color. Once on a planet, the participant must select one of two aliens. Each alien may produce a reward (gemstones) upon being selected, or it may produce rocks (the unrewarded condition). Participants are instructed to maximize their gemstone rewards, forcing participants to optimize their strategy such that they pick the most favorable rocket-ship/alien combination the majority of the time and thus get the highest total number of rewards at the end of the game. Throughout the game, the reward probabilities associated with given aliens vary (via a small Gaussian walk) to prevent memorization and encourage participants to continue to reassess their model of the environment and re-evaluate their optimal strategy. While participants complete the task, the task host site tracks their keyboard inputs and stores them locally in the browser. At the end of a given trial, user metrics are exported and emailed to us for analysis. Metrics include rocket ship and alien choices, whether or not a task run was rewarded, and numerous other related metrics. A full list of collected metrics can be found in the data folder of our project repository.[4]

Participants had to then complete the tiredness assessment and two-step task again at a different time of day (based on their experimental condition). Note that the experimental condition time windows were chosen based on scientifically established times of high and low fatigue (on average) (Valdez, 2019).

Modeling

To model the results of our experiment, we used R and the "lme4" package to perform logistic mixed-effect modeling. As mentioned in the approach section, the use of model-based versus model-free reinforcement learning strategies impacts whether or not a participant chooses to select the same first-stage choice between any given pair of successive task runs (Douglas Bates, 2015). As such, we coded whether or not a participant "stayed" with the same first-stage choice between task runs as a binary variable. This enabled us to predict

[1] https://forms.gle/gqvDWck4GKZFY9Qv6

[2] https://c-loftus.github.io/CognitiveModeling/

[3] https://github.com/nivlab/jspsych-demos/tree/main/tasks/two-step

[4] https://github.com/C-Loftus/CognitiveModeling/tree/master/data

"stay" using logistic mixed-effect modeling with a variety of different predictors. Specifically, we considered whether or not a participant was rewarded in the previous task run ("outcome") and whether the first-to-second stage transition in the previous task run was common (70% probability) or rare (30% probability; "transition"), as these are the necessary predictors required to determine reinforcement strategy usage. We then created subsequent models involving additional predictors, specifically a given participant's quantitative tiredness rating for a given trial, whether trial data came from a participant's first or second trial, and whether the trial data came from the morning or evening.

Our modeling uses the "outcome" coefficient as a proxy for model-free reinforcement learning. This is a reasonable proxy to use, as an entirely model-free learner determines whether or not to stay with the same first-stage choice based on whether or not that first-stage choice resulted in a reward during the previous task run. Our modeling uses the "outcome:transition" interaction coefficient as a proxy for model-based reinforcement learning. This is a reasonable proxy to use, as a model-based learner would analyze both the outcome and the transition taken during the previous task run when determining whether or not to stay with the same first-stage choice (i.e., they would select the same first-stage choice only if a common transition occurred and was rewarded or if a rare transition occurred that was unrewarded). Similar proxies have been used in past iterations of the two-step task (N. D. Daw et al., 2011). Also, to ensure the desired effects of our predictors on stay were statistically reliable at the population level, we made each of our mixed-effects models hierarchical, with all effects taken as random effects across subjects.

Results

Upon running the logistic mixed-effects model to predict stay based on outcome and transition, we found statistically significant effects of both outcome and the outcome-transition interaction on stay. Analyzing results on the population level, the main effect of outcome was statistically significant ($p \approx 0.001$), indicative of a model-free effect. Simultaneously, the interaction between outcome and transition was also significant ($p \approx 0.004$), indicative of a model-based effect. As such, participants incorporate both model-based and model-free strategies into their decision-making. As both reinforcement learning paradigms predict, there was no significant main effect of transition ($p \approx 0.2$). The intercept term was significantly positive ($p < 3e-11$), which indicates a tendency to select the same first-stage choice across task runs, regardless of the outcome. This is consistent with past implementations of the two-step task (N. D. Daw et al., 2011), indicating the correctness of our experimental design and establishing our ability to isolate the effects of fatigue. These results are summarized in Table 1 and Figure 2.

Table 1: Two-Step Task Betas (Fatigue not Included)

	Estimate	Std. Error	$\Pr(>\|z\|)$
Intercept	1.73	0.26	3.01e-11***
outcome	0.13	0.04	0.001**
transition	-0.11	0.09	0.197
outcome:transition	0.24	0.09	0.004**

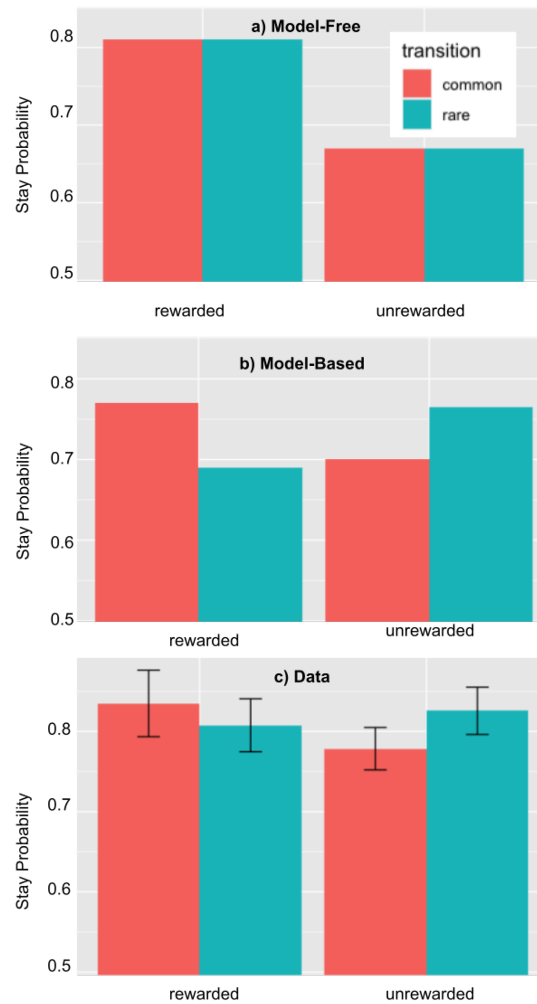

Figure 2: Mixed Effects-Based Analysis of Stay Behavior (A) According to model-free reinforcement learning, first-stage choices resulting in a reward are more likely to be selected on subsequent task runs, regardless of transition type. (B) According to model-based reinforcement learning, the transition type impacts the value of the other first-stage option, causing an interaction between outcome and transition. (C) Experimental stay proportions, averaged across all subjects. Proportions display the hallmarks of both learning strategies. Error bars: 1 SEM.

Factoring in the trial number (i.e., first or second time completing the task) and time of trial (i.e., morning or evening), and again analyzing results on the population level, we did not find a significant relationship between the time of trial and the outcome or between the time of trial and the outcome-transition interaction ($p \approx 0.4$ and $p \approx 0.8$, respectively). We did find the interaction between the trial number and the task outcome statistically significant ($p \approx 0.015$), with second trials exhibiting decreased model-free effects (decreased reliance on solely reward). These results are summarized in Table 2.

Table 2: Two-Step Task Betas (with time of day manipulation)

	Estimate	Std. Error	$\Pr(>\|z\|)$
outcome:trial	-0.14	0.06	0.015*
outcome:morning	-0.04	0.05	0.407
outcome:transition:trial	0.11	0.13	0.391
outcome:transition:morning	-0.03	0.11	0.809

Factoring in self-reported fatigue, and again analyzing results on the population level, we found a significant fatigue-outcome interaction ($p \approx 2e-6$), with decreased fatigue associated with decreased model-free effects (decreased reliance on solely reward). We also found a significant effect of fatigue on the outcome-transition interaction ($p \approx 0.001$), with decreased fatigue associated with decreased model-based effects (decreased reliance on the reward-transition interaction). These results are summarized in Table 3 and Figure 3.

Table 3: Two-Step Task Betas (with fatigue factor)

	Estimate	Std. Error	$\Pr(>\|z\|)$
outcome:fatigue	0.14	0.03	2e-6***
outcome:transition:fatigue	0.20	0.06	0.001**

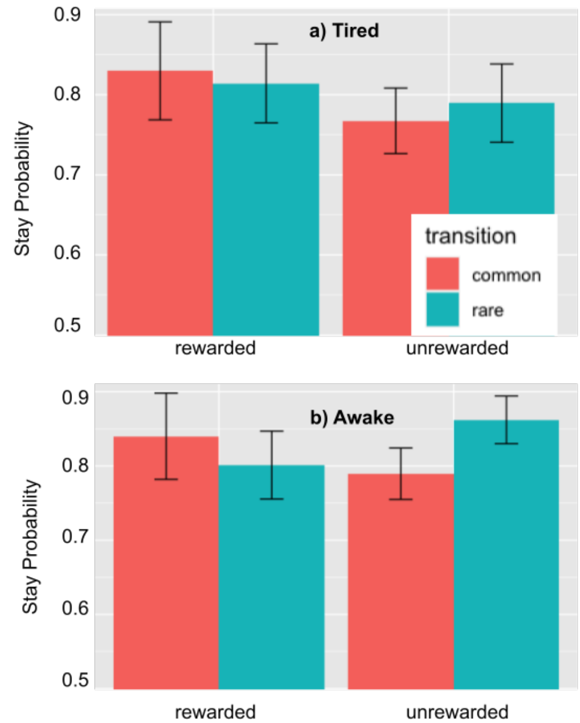

Figure 3: Stay Behavior based on Median-Split Fatigue
(A) Experimental stay proportions, averaged across all trials for which the subject completing the trial had a fatigue level in the lower half of possible fatigue values. Proportions display the hallmarks of both model-based and model-free strategies. Error bars: 1 SEM.
(B) Experimental stay proportions, averaged across all trials for which the subject completing the trial had a fatigue level in the upper half of possible fatigue values. Proportions do not significantly correspond to model-based or model-free strategies. Error bars: 1 SEM.

Discussion

This paper aimed to determine the influence of fatigue on approaches to decision-making tasks. Specifically, we aimed to assess fatigue's influence on people's use of model-based and model-free reinforcement learning during a sequential choice ("two-step") task. Through our analyses, we found that decreased fatigue reduces the usage of both model-based and model-free learning strategies. More broadly, this finding indicates that approaches to decision-making tasks may be influenced by cognitive resource limitation fluctuations across individuals, even when accounting for human cognitive limitations more generally.

Baseline Modeling

Isolating the model-based and model-free effects of fatigue levels, we found results that paralleled two-step task implementations in past literature (N. D. Daw et al., 2011). This establishes the integrity of our experimental setup and enables us to accurately isolate the influence of fatigue.

Influence of Time of Day

The time of day at which participants performed the two-step task (i.e., midday versus night) did not statistically significantly influence their behavior. This result is reasonably intuitive, given that some of our participants were more fatigued during the day, while others were more fatigued at night. Differing times of peak fatigue are expected given our participant demographics – a mixture of adults and college students. These groups likely follow significantly different sleeping schedules and, thus, also have differing levels of fatigue at the same time of day.

Influence of Fatigue

Fatigue statistically significantly impacts the use of reinforcement learning strategies. Specifically, increased alertness (decreased fatigue) reduces the significance of both model-based *and* model-free approaches. This is an unexpected finding, as we initially hypothesized that, under conditions of low fatigue, participants would be increasingly model-based and decreasingly model-free (given that we associated increased alertness with increased access to the cognitive resources necessary to model a task).

Perhaps such a result occurred because subjects with higher energy levels felt more motivated to get the task done quickly, and did not take sufficient time to process the experimental paradigm. The fact that success on our task only resulted in an artificial, digital reward as opposed to a physical, monetary reward may have also shifted subjects' motivation from maximizing reward to completing the task quickly (thus reducing attentiveness to the task). In addition, alert subjects may have been distracted from our task by their "real-life" tasks and responsibilities, given that they had to interrupt the likely most productive hours of the day to complete the "midday" portion of our experiment. Thus, even though not fatigued, these subjects' minds may have been preoccupied with their primary responsibilities and stresses rather than our experiment (i.e., stress and attentiveness may be covariates of fatigue that influenced our experimental findings).

Limitations and Future Work

Future work may benefit from a larger and even more demographically diverse sample, given that this study primarily sampled college students. A larger sample may elucidate whether our unexpected results regarding fatigue are simply due to the aforementioned stress explanation (i.e., people rush to get the task done when it is scheduled during their most productive working hours). Future research in this domain could also consider using an objective measure of fatigue (e.g., a reaction time test) since people might be misreporting their sleepiness (either intentionally or due to a lack of mental awareness). Studying the effects of more extreme fatigue variations (for example, examining subjects before and after ingesting caffeine and/or undergoing sleep deprivation) may also be illuminating. Such stark fatigue variations (as opposed to the mild variations examined in this paper) may result in more significant effects on learning strategies.

Acknowledgements

We would like to thank our graduate student mentor, Evan Russek, and our Professor, Tom Griffiths, for their advice throughout the process of writing this paper and conducting our experiment. We would like to thank our classmate, Gili Karni for answering general questions regarding the NivLab two-step codebase.

In this project, Colton focused primarily on setting up the experiment and website, as well as writing code to automatically process the CSVs. Maya focused primarily on modeling the experimental data and interpreting the results. Both Colton and Maya worked on writing this paper and finding subjects for the study.

References

Almonroeder, T. G., Tighe, S. M., Miller, T. M., & Lanning, C. R. (2018). The influence of fatigue on decision-making in athletes: a systematic review. *Sports Biomechanics*, *19*, 76 - 89.

Daw, N., Niv, Y., & Dayan, P. (2005, 12). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature neuroscience*, 1704–1711. doi: 10.1038/nn1560

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215. doi: 10.1016/j.neuron.2011.02.027

Dezfouli, A., & Balleine, B. W. (2013, 12). Actions, action sequences and habits: Evidence that goal-directed and habitual action control are hierarchically organized. *PLOS Computational Biology*, *9*(12), 1-14. Retrieved from `https://doi.org/10.1371/journal.pcbi.1003364` doi: 10.1371/journal.pcbi.1003364

Douglas Bates, B. B. S. W., Martin Machler. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: 10.18637/jss.v067.i01

Griffiths, T., Lieder, F., & Goodman, N. (2015). Rational use of cognitive resources: levels of analysis between the computational and the algorithmic. *Topics in cognitive science*, *7*(2), 217–229. doi: 10.1111/tops.12142

Khalvati, K., Park, S. A., Mirbagheri, S., Philippe, R., Sestito, M., Dreher, J.-C., & Rao, R. P. N. (2019). Modeling other minds: Bayesian inference explains human choices in group decision-making. *Science Advances*, *5*(11), eaax8783. Retrieved from `https://www.science.org/doi/abs/10.1126/sciadv.aax8783` doi: 10.1126/sciadv.aax8783

Mullette-Gillman, O. A., Leong, R. L. F., & Kurnianingsih, Y. A. (2015). Cognitive fatigue destabilizes economic decision making preferences and strategies. *PLoS ONE*, *10*.

Persson, E., Barrafrem, K., Meunier, A., & Tinghög, G. (2019). The effect of decision fatigue on surgeons' clinical decision making. *Health Economics*, *28*, 1194 - 1203.

Shahid, A., Wilkinson, K., Marcu, S., & Shapiro, C. (2011, 11). Stanford sleepiness scale (sss). *STOP, THAT and One Hundred Other Sleep Scales*, 369-370. doi: 10.1007/978-1-4419-9893-4_91

Valdez, P. (2019). Circadian rhythms in attention. *The Yale journal of biology and medicine*, *92*(1), 81-92.

Wiehler, A., Branzoli, F., Adanyeguh, I., Mochel, F., & Pessiglione, M. (2022, August 22). A neuro-metabolic account of why daylong cognitive work alters the control of economic decisions. *Current Biology*, *32*(16), 3564–3575.e5. doi: 10.1016/j.cub.2022.07.010