

The Effects of Mental Fatigue on Effort Allocation: Modeling and Estimation

Zhide Wang¹, Yanling Chang², Brandon J. Schmeichel³, and Alfredo A. Garcia¹

¹ Department of Industrial and Systems Engineering, Texas A&M University

² Department of Engineering Technology and Industrial Distribution, Department of Industrial and Systems Engineering,
Texas A&M University

³ Department of Psychological and Brain Sciences, Texas A&M University

Mental fatigue is usually accompanied by drops in task performance and reduced willingness for further exertion. A value-based theoretical account may help to explain such negative effects. In this view, mental fatigue influences perceived costs and rewards of exerting effort. However, no formal mathematical framework has yet been proposed to model and quantitatively estimate the effects of mental fatigue on subjective evaluations of effort expenditure, under possibly imperfect self-perceptions of internal fatigue states. We proposed a mathematical framework to model human cognitive effort allocations, assuming mental fatigue states are partially observable with semi-Markov dynamics. We modeled effort allocation decisions as a means to the goal of maximizing cumulative subjective values over a given time horizon. We developed an estimation method to identify subjective values and the hidden dynamics of mental fatigue, which can in future work be applied to self-reports, psychophysiological indices, and behavioral outcomes associated with fatigue. The modeling and estimation method was tested using a simulated *n*-back task under a free-choice paradigm, with model parameters fine-tuned from past studies. The proposed approach was able to recapitulate task performance and task engagement patterns observed under mental fatigue. This work advances a reward/cost trade-off account for explaining the exertion of mental effort and suggests new avenues for both theoretically and empirically relevant understandings of how cognitive operations are affected by mental fatigue.

Keywords: mental fatigue, effort allocation, decision-making, estimation

Supplemental materials: <https://doi.org/10.1037/rev0000365.supp>

Mental fatigue is associated with reduced effort on demanding cognitive operations. Consequently, task performance and work efficiency tend to decline under mental fatigue (Boksem et al., 2005, 2006; Hopstaken et al., 2015, 2016). The adverse consequences of mental fatigue have been documented in a variety of settings. In an office work environment, for example, mental fatigue has been observed to reduce both typewriting speed and accuracy (de Jong et al., 2020). In transportation, mental fatigue among drivers has been considered a major cause of road accidents (Lal & Craig, 2001). It has been estimated that 20%–30% of all road accidents and 5%–15% of fatal road accidents involve fatigue (Razmpa et al., 2011). In health care, more than 78% of health providers including physicians and nurses have reported experiencing fatigue-related burnout, which has been identified as “a public health crisis that urgently demands action” (Jha et al., 2019).

Although mental fatigue has been studied formally for over 100 years, and its adverse consequences have been observed for even longer (Thorndike, 1900), its origin and effects are still a matter of

debate (Massar et al., 2018; Pattyn et al., 2018). In this article, we propose a mathematical framework for modeling and estimating the effects of mental fatigue on the allocation of mental effort. The guiding assumption in this approach is that humans are rational in deciding to exert mental effort in the sense that they aim to maximize subjective values (Loewenstein et al., 2008; Shenhav et al., 2013; Westbrook & Braver, 2015). Simply put, when the rewards for successful task performance exceed the perceived costs, people are inclined to put effort into the task. Otherwise, the task may be abandoned in favor of other more valued activities. We assume that mental fatigue impacts the subjective value of effort expenditure, and we report a mathematical model rooted in this perspective that may help to estimate and explain the effects of mental fatigue.

Mental Effort

Many cognitive activities are habitual and can be executed quickly and easily, but some mental operations—especially those

Zhide Wang  <https://orcid.org/0000-0002-6970-399X>
Yanling Chang  <https://orcid.org/0000-0002-0935-5735>
Alfredo A. Garcia  <https://orcid.org/0000-0002-2761-7479>

The present study was not preregistered and there was no prior

dissemination of the ideas presented in the present article.

Correspondence concerning this article should be addressed to Zhide Wang, Department of Industrial and Systems Engineering, Texas A&M University, 400 Bizzell Street, College Station, TX 77843, United States. Email: liang93429@tamu.edu

involving cognitive control—are slower and more effortful and thus may be construed as costly. For example, encoding a single-digit number and recalling it a few moments later is a simple cognitive operation that even young children, who lack fully developed cognitive capacities, can accomplish with relative ease. By contrast, encoding a five-digit number while processing a dense passage of text and then recalling the number minutes later is a more effortful mental operation. Juggling multiple streams of information requires cognitive control to maintain the target information (i.e., the number) in working memory and to prevent it from being dislodged by competing information (Engle, 2010; Unsworth & Engle, 2007). But what currency is used to pay for costly mental effort?

The costs of cognitive operations have at least two origins, namely, intrinsic costs and opportunity costs. Intrinsic costs stem from the limited nature of cognitive capacity: only so much information can be processed at any given time. Such intrinsic costs may derive from limited metabolic resources or finite task-relevant information storage capacity (Shenhav et al., 2017). Indeed, one stream of thought has suggested that suboptimal behaviors are deeply rooted in the finite nature of the brain's computational capacity (Gershman et al., 2015; Griffiths et al., 2015). Opportunity costs arise insofar as more than one task is available to perform. Exerting mental effort on one task comes at the price of foregoing other tasks and the rewards associated with them. Most relevant to this article is the assumption that the way people allocate mental effort reflects their subjective preferences and optimality principles (Lieder et al., 2012; Vul et al., 2014).

The limits of cognitive capacity—whether due to intrinsic costs or opportunity costs—call for deliberate allocation of mental effort. Numerous experiments have observed that individuals adjust the level of effort they allocate to cognitive tasks according to the perceived trade-offs between task-associated costs and rewards. For example, when allowed to choose among different tasks to perform, participants avoided more cognitively demanding tasks in favor of less demanding tasks (Kool et al., 2010). Another study used an economic discounting paradigm to study subjective costs associated with different tasks (Westbrook et al., 2013). Participants were instructed to choose among a set of tasks associated with different difficulty levels and different rewards. The rewards required for choosing to perform highly demanding tasks were substantially higher than those for simpler tasks. These results revealed a reluctance to exert effort when available rewards for successful task performance fall below a subjective break-even point.

The Subjective Value of Effort Expenditure

The difference between the perceived costs and rewards for performing a task can be construed as stemming from subjective value (Westbrook & Braver, 2015). Building upon the concept of subjective value, a normative model can be formulated to translate a person's behaviors into subjective preferences (and vice versa). Moreover, the model parameters can, in some cases, be estimated from behavioral data (Chong et al., 2017; Hartmann et al., 2013; Klein-Flügge et al., 2015; Prévost et al., 2010). For example, one investigation attempted to fit linear, hyperbolic, parabolic, and exponential subjective value functions to participants' choices to allocate mental and physical effort by using a maximum likelihood estimation procedure (Chong et al., 2017). This work yielded different models for mental (hyperbolic) versus physical effort

(parabolic), a conclusion further corroborated by evidence for both domain-general and domain-specific neural substrates for coding the two types of effort.

An important strand in the literature on modeling mental exertion by subjective values relates to the expected value of control (EVC) model (Shenhav et al., 2013, 2017), which construes the subjective value of effortful cognitive control as the sum of the value of the present state of the world and the expected future value that results from the current effort expenditures. According to the EVC model, the allocation of mental effort reflects a rational decision process whereby effort is adjusted over time to maximize subjective value. The EVC model extended a value-based decision-making account by considering the consequences of current cognitive effort exertion on distal subjective values (Chong et al., 2017; Hartmann et al., 2013; Klein-Flügge et al., 2015; Prévost et al., 2010).

The EVC model has been tested and validated in a handful of prior studies. For example, one study built an EVC model using a simulated flanker task (Musslick et al., 2015). This study validated classic findings in a sample of simulated subjects, including reduced error rates and slower reactions after erroneous responses (Laming, 1968), as well increased intensity of mental effort with increased incentives for task performance (Padmala & Pessoa, 2010). The EVC model also has provided a useful framework for studying the plasticity of control (Lieder et al., 2018) and the influence of affect on cognitive control (Grahek et al., 2020).

However, some basic components of the EVC model are not directly observable and thus are hard to identify. For example, the value of the outcomes and the costs of cognitive control cannot be directly observed and thus the EVC model might be mis-specified, which limits the capability of the EVC model to explain empirical data. Hence, to validate the EVC framework, those model components have to be correctly identified. Although some past work has been done to obtain those basic components from experimental data (Musslick et al., 2018), there is still far from enough discussion on how best to identify otherwise unknown model primitives. Nevertheless, the EVC model, as we shall propose below, can be extended to incorporate and identify the effects of mental fatigue.

Effects of Mental Fatigue on Cognitive Control

There is no gold standard definition of mental fatigue (Massar et al., 2018; Pattyn et al., 2018). As mentioned previously, both common experience and empirical studies have revealed that mental fatigue is accompanied by aversion to effort expenditure, poorer task performance, and subjective feelings of tiredness (Boksem et al., 2005; Dittner et al., 2004; Hopstaken et al., 2015, 2016; Pattyn et al., 2018). A prevailing metaphor for mental fatigue, similar to physical fatigue, is that it is induced by the depletion of energetic resources. The energetic resource has been hypothesized to be glucose in the bloodstream, which is essential for virtually all bodily (including brain) functions (Fairclough & Houston, 2004; Gailliot et al., 2007). The idea is that the resource (e.g., glucose) gets used up by performing effortful mental tasks, leaving the person relatively less able to continue to exert effort until the resource has been restored. However, the resource metaphor has been challenged both conceptually and empirically. Whereas some studies found evidence that blood glucose decreases after cognitively demanding tasks (Fairclough & Houston, 2004; Gailliot & Baumeister, 2007), others found no global metabolic changes after mental effort expenditure

(Finley et al., 2019; Westbrook & Braver, 2015). Blood glucose levels do not vary much if at all during cognitive demanding tasks (Gibson, 2007; Kurzban, 2010). Even “resting” mental activity could be construed as metabolically costly, insofar as specific task performance hardly consumes more brain energy than does resting activity (Raichle & Mintun, 2006).

As an alternative to the resource-based view, some theorists have endorsed a more motivationally oriented perspective of mental fatigue (Hockey, 2011). In this view, mental fatigue alters the willingness (rather than the ability) to expend effort. This perspective has received support from studies testing the effects of external rewards on mental effort. Both monetary incentives (Boksem et al., 2006; Muraven & Slessareva, 2003) and nonmonetary incentives (Hopstaken et al., 2015, 2016) have been found to reverse the characteristic signs of task disengagement associated with mental fatigue. Hence, even after fatigue has set in, performance and associated physiological indicators of attention control can be restored to (or near to) pre-fatigue levels by offering incentives for performance. These results supported the idea that individuals tend to invest precious cognitive capacities in tasks that yield higher rewards (compared to the costs), and increasing the value of a task (i.e., increasing motivation) makes expending effort on it more likely. In terms of subjective value, we propose that mental fatigue affects the perceived costs and rewards of mental operations and, consequently, the motivation for exerting mental effort. Changing the reward/cost balance by introducing incentives for performance can thus modulate the allocation of effort under mental fatigue.

A sound model of mental fatigue should be able to explain some crucial established empirical findings, including evidence that (a) mentally fatigued people become inclined to disengage from effortful tasks and reduce further effort exertion (Giacomantonio et al., 2014, 2019; Hopstaken et al., 2016) and (b) extra incentives help to offset or reverse these negative effects of mental fatigue (Boksem et al., 2005; Hopstaken et al., 2015). Below we develop a mathematical model intended to account for these established patterns. We intend the proposed approach to be used in future work to further understand the causes and consequences of mental fatigue.

Contribution

To our knowledge, the existing literature lacks modeling and estimation methods for testing the extent to which fatigue changes during cognitive exertion, and the extent to which perceived rewards and costs of mental effort change under mental fatigue. In this article, we developed a quantitative model and estimation methods to examine how mental fatigue affects the allocation of effort during cognitively demanding tasks.

In the proposed model, at each decision epoch, a human subject (agent) decides the amount of effort to exert, expecting performance-related rewards in return for the costs of mobilizing mental effort. We assume that the agent chooses the amount of mental effort expenditure that maximizes total subjective value over the extended task duration, and further we assume that the fatigue state is imperfectly self-perceived by the agent. The transition between fatigue states follows a hidden semi-Markov process, that is, the likelihood of transition to a more fatigued state is a function of the current state and the elapsed time in the current state. We take a “Bayesian brain” perspective, which postulates that the agent infers hidden causes of sensations based on generative models

(Friston, 2012; Knill & Pouget, 2004),¹ and assume that humans represent sensory information of hidden fatigue states through Bayesian belief and update the belief upon encountering new input data (a Bayesian belief is the posterior distribution over possible fatigue states, given sensory related and behavioral data).²

We develop a method to estimate the parameters of the model, including the hidden fatigue dynamics and the subjective values of effort expenditure, using sensory and behavioral data as well as effort allocation decisions. We tested the estimation method via a simulated n -back task under a free-choice paradigm with model parameters fine-tuned from the published literature. The aims were to identify the model primitives and to correctly predict common empirical findings such as the decline in performance and task engagement due to increased fatigue (Giacomantonio et al., 2019; Hopstaken et al., 2016) and task reengagement due to extra incentives (Giacomantonio et al., 2014; Giacomantonio et al., 2019; Hopstaken et al., 2015). This work advances methods for explaining effort allocation decisions and provides new avenues for empirically relevant understanding of how effort expenditure is affected by mental fatigue.

Review of Relevant Mathematical Background

Partially Observable Markov Decision Processes

Markov decision processes (MDPs) are used to model optimal intertemporal decision-making (Bellman, 1966; Puterman, 2014; Ross, 2014). At each decision epoch $t > 0$, the agent observes the state of the environment s_t and selects an action (makes a decision) a_t . The action a_t yields a reward $r(s_t, a_t)$ (depending on both state and action) and leads to a change in the state of the environment. The transition of the state is characterized in a probabilistic manner by a conditional probability $P(s_{t+1}|s_t, a_t)$. The agent observes s_{t+1} before making the next choice and so on and so forth.

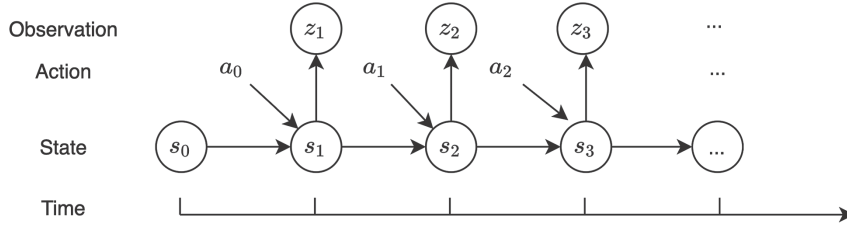
Partially observable Markov decision processes (POMDPs) generalize MDPs by taking into account the fact that the states of the environment may not be perfectly observed (Cassandra, 1994; Monahan, 1982; Smallwood & Sondik, 1973).³ Due to the hidden nature of environment states, the agent has to infer the state and makes decisions based upon all possible observed histories (including all past decisions and state-correlated noisy measurements); see Figure 1. The consequence is that the relevant history grows exponentially as more decision epochs are considered. Fortunately, for a Bayesian agent, a belief—a probability distribution over the state space—is sufficient to derive an optimal policy. The decision-making process is similar to MDPs: At decision epoch t , the agent observes a hidden state-correlated observation z_t and updates belief x_t based on previous belief x_{t-1} , previous action a_{t-1} and z_t . A decision a_t is made based on x_t . Upon execution of the decision a_t , the agent receives a reward and its expectation is denoted by

¹ In the context of a Bayesian brain, a generative model refers to the modeling approach that describes hidden causes X and their dependent observables Y via a joint distribution $P(X, Y)$ (Jebara, 2012; Ng & Jordan, 2002).

² Sensory related data may include electroencephalography, electrocardiography, pupil diameter, etc. Behavioral data may include task performance indices such as reaction times, error rates, etc.

³ In general, the environment includes both external states such as temperature and humidity, and internal states. In this work, by environment states we mean human psychological states, or more specifically, mental fatigue states.

Figure 1
Structure of POMDPs



Note. POMDPs = Partially observable Markov decision processes. At each decision epoch t , a hidden environment state s_t emits a correlated observation z_t . The agent makes a decision by consulting z_t in addition to all past actions and observations. The hidden state s_t is influenced by the selected action a_t and transitions to s_{t+1} .

$r(x, a)$. A transition of the hidden state resulting from decision a_t is described by $P(s_{t+1}, z_{t+1} | s_t, a_t)$. The goal for the agent is to maximize the expected cumulative reward: $E[\sum_{t=0}^{\infty} \beta^t r(s_t, a_t) | x_0]$. A six tuple $(\mathbb{S}, \mathbb{A}, \mathbb{Z}, P, r, \beta)$ formally defines a POMDP, where \mathbb{S} is the set of possible hidden states, \mathbb{A} is the set of available actions, \mathbb{Z} is the set of possible observations, the dynamics of the hidden state is $P: \mathbb{S} \times \mathbb{A} \rightarrow \mathbb{S} \times \mathbb{Z}$, and $r: \mathbb{S} \times \mathbb{A} \rightarrow \mathbb{R}$ is the reward function (\mathbb{R} is the space of real numbers). MDPs and POMDPs have been used in modeling human behaviors—such as goal-directed decision-making under uncertainty—and their biological foundations (Dayan & Daw, 2008; Seymour et al., 2004).

Inverse Reinforcement Learning and Dynamic Discrete Choice Models

Inverse reinforcement learning (IRL) is an approach to estimate the preferences (modeled by reward functions) of an agent in goal-directed behavior, in order to explain why some behaviors are more preferred than others (Abbeel & Ng, 2004; Arora & Doshi, 2018; Dvijotham & Todorov, 2010; Ng & Russell, 2000). Note that although both IRL and MDP approaches attempt to model optimal behavior in the sense of utility maximization, the two approaches are different: The goal of MDPs is to derive an optimal policy for a given (known) set of model parameters, whereas IRL assumes the optimal behaviors are exhibited in the data but the corresponding model parameters are unknown. The goal of IRL is to “learn” the parameters from the given data (hence “inverse”).

Specifically, for a classic model-based IRL approach, the goal is to learn the reward function (i.e., the agent’s relative preferences for the available options). Whereas IRL may be capable of learning human driving behaviors (Kudrner et al., 2015), training robots (Finn et al., 2016), and understanding others’ intentions (Theory of Mind; Jara-Ettinger, 2019), the method has shortcomings. One crucial issue overlooked by the IRL approach stems from its assumption that the agent perfectly observes the environment. This assumption is almost certainly inaccurate. Even if the environment were fully observable, human perceptions and measures are noisy and ambiguous (Schmitt et al., 2017). One study attempted to tackle this issue by developing a model to estimate the reward function in a partially observable environment (Choi & Kim, 2011). However, in that work, how the state of the environment changes

(i.e., environment state dynamics) was assumed to be known. This assumption was likely overly restrictive and simplistic: Changes in many subjective states (such as mental fatigue) are context dependent and influenced by individual differences and hence may not be known in advance.

Econometricians have also developed estimation methods for inferring the parameters of the agent’s utility model from decision data, which are generally denoted by *structural estimation*. Rust (1987) authored a seminal work that extended the scope of structural estimation to MDPs by developing a nested fixed point method to estimate (completely observed) environment state dynamics and reward function, based on an agent’s decision-making history. Following Rust (1987), a class of dynamic discrete choice (DDC) models was developed to describe (and estimate and predict) the decision-making behaviors of a forward-looking agent (Aguirregabiria & Mira, 2010; Keane et al., 2011). DDC models have been used in the study of human economic behaviors such as labor supply (Heckman, 1974), education policy (Todd & Wolpin, 2006), career decisions among politicians (Diermeier et al., 2005), and many others. However, most DDC models assume that the environment is fully observable. The limitations associated with assuming fully observable states have recently garnered attention. Chang et al. (2020) generalized the Rust (1987) model to incorporate hidden states with Markov dynamics, and a maximum likelihood estimation procedure was developed to recover both the reward function and the hidden state dynamics. In the present work, we will generalize Chang et al.’s method to include semi-Markov states and apply it to learning the hidden dynamics of mental fatigue and the subjective values of a human agent in the context of mental effort allocation.

Hidden Markov Models and Hidden Semi-Markov Models

The class of hidden Markov models (HMMs) is a statistical framework for the Markov processes in which the environment states are unobserved. Similar to POMDPs, an observable process (observation) is available in an HMM and is dependent on the underlying hidden Markov process. The difference between HMMs and POMDPs is that HMMs do not involve actions (i.e., no decision-making process). At time $t > 0$, the environment

occupies a (hidden) state s_t and the environment releases an observable signal z_t dependent on s_t . Moreover, the underlying s_t transitions to s_{t+1} stochastically, following the hidden state dynamics $P(s_{t+1}|s_t)$; see Figure 2. The three major goals of HMMs are: (a) to compute the likelihood of an observed data sequence $\{z_t\}$, (b) to find the most likely hidden state sequence $\{s_t\}$ for a given history of observations, and (c) to learn hidden state dynamics that define an HMM (Baum & Petrie, 1966; Baum et al., 1970; Eddy, 2004; Mor et al., 2021).

Sojourn time describes the amount of time one spends in an environment state before exiting that state. In the study of mental fatigue, sojourn time may be used to characterize the amount of time an individual remains in a particular fatigue state (e.g., fatigued or not fatigued). Classic HMMs are restricted to the case wherein the sojourn time distribution of a hidden state is memoryless. The class of hidden semi-Markov models (HSMMs), a variation of HMMs, allows for more flexible state-duration (sojourn time) modeling (Yu, 2010). See Figure 3 for an illustration of the structure of HSMMs. In an HSMM, when the environment transitions to a different state, it remains in that state for d time periods and emits d (possibly different) observations. The distribution of duration d can follow a variety of probability distributions (Mor et al., 2021). In other words, HSMMs explicitly consider and model the time the environment has spent in the current state s , or the time left in the current state s (before transitioning to another state $s' \neq s$). HMMs and HSMMs have enabled wide applications in the study of mental fatigue (Fu et al., 2016; Liu et al., 2010; Luo & Qi, 2019; Zhang et al., 2008), and other psychological states such as affect (D'Mello & Graesser, 2010; Le & Provost, 2013), and brain electrical activity (Doroshenko et al., 2007; Obermaier et al., 2001; Power et al., 2010; Wang et al., 2013).

A Mathematical Framework of Effort Allocation Under Mental Fatigue

In this section, we propose a model to formalize the effects of mental fatigue on effort allocation. We take a reward/cost trade-off perspective, attributing mental effort allocation to subjective values.⁴ In this section, we focus on modeling mental effort allocation as the maximization of the expected cumulative subjective values over the duration of a cognitively demanding task. A concrete example of mental effort allocation including a parameterization of costs and rewards under the effects of mental fatigue will be described below (under N -back Task).

At each decision epoch $t > 0$, the agent's fatigue state $s_t \in \mathbb{S}$ is not directly observable (either to the agent or to observers), although a noisy observation $z_t \in \mathbb{Z}$ (e.g., a physiological or behavioral measure) of the fatigue state is available. Let $a_t \in \mathbb{A}$ be the effort allocation choice and $\tau_t \in \mathbb{D}$ be the duration of the relevant inner state (hidden to both the agent and to observers), indicating the time left in the current fatigue state at time t . In this work, the fatigue state space \mathbb{S} , the observation space \mathbb{Z} , the action space \mathbb{A} , and the temporal duration space \mathbb{D} are assumed to be finite. The intertemporal transition of both fatigue state and its duration is denoted by transition probabilities $\{P(s', \tau', z|s, \tau, a) : s, s' \in \mathbb{S}, z \in \mathbb{Z}, \tau', \tau \in \mathbb{D}, a \in \mathbb{A}\}$.⁵

One-step subjective value $r(s, a)$ is the trade-off between rewards and costs for a specific decision epoch, which is a function of both fatigue state s and selected effort level a .

Subjective value functions could be nonparametric or parametric (such as linear, parabolic, hyperbolic, or exponential functions of both fatigue level s and effort level a). It is possible that z_t may not fully capture the information that the agent has available for decision-making (e.g., the agent may possess private information that influences subjective values and decisions). We thus assume that one-step subjective value is perturbed by a random factor $\varepsilon = \{\varepsilon(a) : a \in \mathbb{A}\}$, wherein the a th element of vector ε stands for the privately observed perturbation when the agent chooses to execute decision a .⁶ See Figure 4 for a schematic representation.

Let $\zeta_t = \{x_0, a_0, z_1, a_1, z_2, \dots, a_{t-1}, z_t\}$ be the history until time t , containing all observable information including past effort allocation choices $\{a_0, a_1, \dots, a_{t-1}\}$ and past observations $\{z_0, z_1, \dots, z_t\}$, wherein $\zeta_0 = \{x_0\}$ and $x_0 = \{P(s_0, \tau_0) : s_0 \in \mathbb{S}, \tau_0 \in \mathbb{D}\}$ is the agent's initial belief (self-perception) of his or her fatigue state and its duration. Assume the agent is rational with the goal of maximizing expected cumulative subjective values over task periods T :

$$\max_{a_0, \dots, a_T} E \left[\sum_{t=0}^T \beta^t (r(s_t, a_t) + \varepsilon(a_t)) | \zeta_0 \right], \quad (1)$$

where E is the expectation operator conditional on ζ_0 and $\beta \in [0, 1)$ is the discount factor.⁷

Following the convention of DDC models, we impose two standard assumptions on random perturbations ε .

Assumption 1 (Conditional Independence).

$$P(z_{t+1}, \varepsilon_{t+1} | \zeta_t, a_t, \varepsilon_t) = P(\varepsilon_{t+1}) P(z_{t+1} | \zeta_t, a_t).$$

Assumption 1 restricts the random perturbation ε_t to be independently and identically distributed (*i.i.d*) over both its own history $\{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots\}$ and other factors (such as fatigue state s and noisy measurements z). Similarly, we assume that the reduced-form observation probability $P(z_{t+1} | \zeta_t, a_t)$ is independent of ε_{t+1} (i.e., we posit that small perturbations in the preference for a task do not dramatically alter task performance or physiological states).⁸

⁴ Subjective values are construed as the difference between potential rewards for successful task performance and the costs incurred to achieve those rewards. Subjective values can be constructed in many ways by using various formulations of rewards and costs (Atkinson, 1957; Chong et al., 2017; Hartmann et al., 2013; Klein-Flügge et al., 2015).

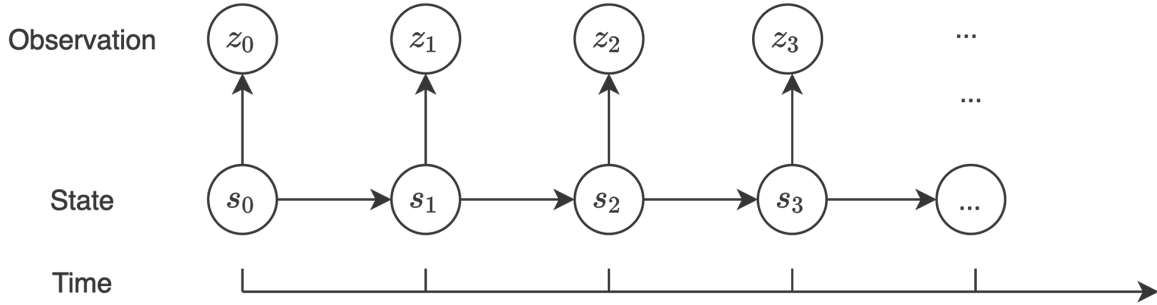
⁵ In reality, a person may experience faster or slower fatigue deterioration depending on his or her current state and the task at hand. As a result, sojourn times for different fatigue states may vary.

⁶ The information discrepancy between agents and observers may contribute to the bounded rationality observed in many human behaviors, which may help to account for why people sometimes appear to behave nonoptimally (Gershman et al., 2015; Griffiths et al., 2015; Puranam et al., 2015; Shenav et al., 2017).

⁷ Some readers may find that the proposed model (1) can be seen as a partially observable semi-Markov decision process (POSMDP) with reward perturbation ε . Whereas the POSMDP literature has focused on computing an optimal control strategy for a known set of model parameters, our goal is to model mental fatigue and estimate the (unknown) parameters of a model based upon observable data.

⁸ *Assumption 1* allows for effective computation of expected cumulative subjective values. However, small perturbations in task preferences may alter task performance or physiological states. Examining how ε may affect physiological states or performance requires a more complicated dependent structure of state dynamics $P(z_{t+1}, \varepsilon_{t+1} | \zeta_t, a_t, \varepsilon_t)$, which represents a potentially valuable avenue for future research.

Figure 2
Structure of HMMs



Note. HMMs = hidden Markov models. At each time t , the environment state s_t is hidden but a state-dependent observation z_t will be emitted. The hidden environment state transitions stochastically, according to a Markov chain (no actions are involved).

Assumption 2 (Additively Separable One-Step Subjective Values).

$$\sum_{s_t, \tau_t} P(s_t, \tau_t | \zeta_t, \varepsilon_t) r(s_t, a_t) = \sum_{s_t, \tau_t} P(s_t, \tau_t | \zeta_t) r(s_t, a_t) + \varepsilon_t(a_t).$$

Assumption 2 states that random perturbation ε contributes to (one-step) subjective values as a random factor, characterizing unexpected or unpredictable influences or bounded rationality when choices are made.

Under *Assumptions 1–2*, the expected cumulative subjective value from time t can be written via the Bellman Equation (Bellman, 1966):

$$g_t(\zeta_t, \varepsilon_t) = \max_{a_t} \left\{ \sum_{s_t} \sum_{\tau_t} P(s_t, \tau_t | \zeta_t) r(s_t, a_t) + \varepsilon_t(a_t) + \beta \sum_{z_{t+1}} P(z_{t+1} | \zeta_t, a_t) \int g_{t+1}(\zeta_{t+1}, \varepsilon_{t+1}) d\mu(\varepsilon_{t+1}) \right\}, \quad (2)$$

where μ is the cumulative distribution function of multivariate random variable ε (of dimension $|\mathbb{A}|$).

We call $g_t(\zeta_t, \varepsilon_t)$ the expected cumulative subjective value function, which can be decomposed into two parts: a present value and an expected future value. From the perspective of the decision-making agent (ε is known before acting):

1. The first part $\sum_{s_t} \sum_{\tau_t} P(s_t, \tau_t | \zeta_t) r(s_t, a_t) + \varepsilon_t(a_t)$ is the perturbed one-step expected subjective value for choice a_t , given the self-perception of both fatigue state and its duration inferred from observable history $P(s_t, \tau_t | \zeta_t)$.
2. The second part $\beta \sum_{z_{t+1}} P(z_{t+1} | \zeta_t, a_t) \int g_{t+1}(\zeta_{t+1}, \varepsilon_{t+1}) d\mu(\varepsilon_{t+1})$ is the forward-looking expected cumulative subjective value function, which weighs every possible future scenario $(z_{t+1}, \varepsilon_{t+1})$ by their probabilities $(P(z_{t+1} | \zeta_t, a_t), d\mu(\varepsilon_{t+1}))$. Specifically, at time $t+1$, different observations (z_{t+1}) may be emitted by the underlying hidden state s_{t+1} , which, together with choice taken at time $t(a_t)$, updates the history ζ_t to ζ_{t+1} (i.e., $\zeta_{t+1} = \{\zeta_t, a_t, z_{t+1}\}$).

Moreover, random perturbation (ε_{t+1}) may also take a different value. So there are two layers of expectation $\left(\sum_{z_{t+1}} P(z_{t+1} | \zeta_t, a_t) \int \dots d\mu(\varepsilon_{t+1}) \right)$, corresponding to the expectation over observations and the expectation over random perturbations.

Finding solution to Equation 2 is difficult as ζ grows exponentially as time increases. To resolve this issue, let \mathbb{X} be the (finite dimensional) probability simplex in $\mathbb{R}^{|\mathbb{S} \times \mathbb{D}|}$ and $x_t \triangleq \{P(s_t, \tau_t | \zeta_t) : (s_t, \tau_t) \in \mathbb{S} \times \mathbb{D}\} \in \mathbb{X}$ be the Bayesian belief (or belief, used interchangeably from here on), which is the posterior distribution over possible fatigue states and its duration (s_t, τ_t) given history ζ_t up to time t . Compared to ζ_t , the size of belief x_t remains constant and does not grow to infinity as t increases (i.e., $x_t \in \mathbb{X}, \forall t$), which enables an efficient solution to Equation 2. According to Bayes rule, upon observing a_t, z_{t+1} , the belief can be updated as:

$$x_{t+1}(s, \tau) = P(s_{t+1} = s, \tau_{t+1} = \tau | z_{t+1}, a_t, \zeta_t) = \frac{\sum_{s_t, \tau_t} P(s_{t+1} = s, \tau_{t+1} = \tau, z_{t+1} | s_t, \tau_t, a_t) x_t(s_t, \tau_t)}{\sigma(x_t, a_t, z_{t+1})}, \quad (3)$$

wherein the denominator is the reduced observation probability:

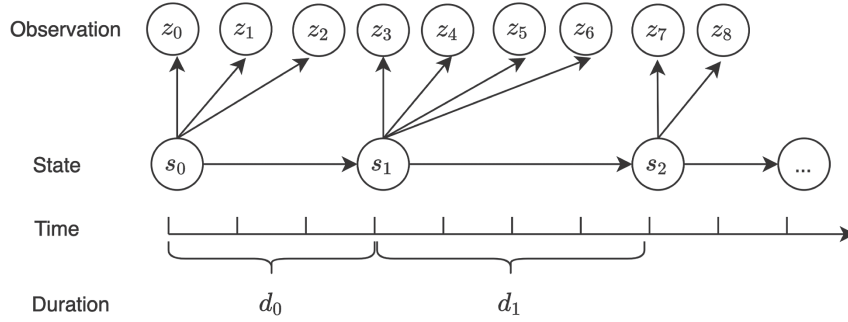
$$\sigma(x_t, a_t, z_{t+1}) \triangleq P(z_{t+1} | x_t, a_t) = \sum_{s_{t+1}, \tau_{t+1}} \sum_{s_t, \tau_t} P(s_{t+1}, \tau_{t+1}, z_{t+1} | s_t, \tau_t, a_t) x_t(s_t, \tau_t). \quad (4)$$

We denote the belief updating process by $x_{t+1} = \lambda(x_t, a_t, z_{t+1})$.⁹ With the definition of belief x , the following theorem formally states that the Bayesian belief x is sufficient to determine the cumulative subjective value function g .

Theorem 1. Let $\zeta_t = \zeta$ be the history of observations and actions up to time t , the belief state, action and random reward perturbation at time t are $x_t = x, a_t = a, \varepsilon_t = \varepsilon$. Define

⁹ More details pertaining to the belief updating process can be found in the Appendices A–C.

Figure 3
Structure of HSMMs



Note. HSMMs = hidden semi-Markov models. The environment remains in a state s for duration d , during which the observation is stochastically dependent on the current occupied state s . The transition of a hidden environment state to a different state is probabilistic once the time left in the current state is 0.

$$V_t(x, \varepsilon) \triangleq \max_a \left\{ r(x, a) + \varepsilon(a) + \beta \sum_{z'} \sigma(x, a, z') \int V_{t+1}(x', \varepsilon') d\mu(\varepsilon') \right\}, \quad (5)$$

where z', ε' are the observation and random perturbation vector at time $t + 1$, the updated belief at $t + 1$ is $x' = \lambda(x, a, z')$ and $r(x, a) = \sum_{s, \tau} r(s, a) x(s, \tau)$. Then

$$V_t(x, \varepsilon) = g_t(\zeta, \varepsilon).$$

Hence, the belief x is a sufficient statistic for the expected cumulative subjective value function g .

Theorem 1 shows that the proposed model in Equation 2 is indeed reformulated to a POMDP-based DDC model in (Chang et al., 2020). Hence, it has the following implications.¹⁰ For infinite T there is a unique (fixed point) value function V such that:¹¹

$$V(x, \varepsilon) = \max_a \left\{ r(x, a) + \varepsilon(a) + \beta \sum_{z'} \sigma(x, a, z') \int V(x', \varepsilon') d\mu(\varepsilon') \right\},$$

where $x' = \lambda(x, a, z')$. The existence of the value function V , in the context of mental effort allocation, enables us to analyze mental effort expenditure behaviors as follows. Let

$$Q(x, a) = r(x, a) + \beta \sum_{z'} \sigma(x, a, z') \int V(x', \varepsilon') d\mu(\varepsilon'), \quad (6)$$

where $x' = \lambda(x, a, z')$. Then the optimal choice a^* for the agent is

$$a^* \in \arg \max_a \left\{ Q(x, a) + \varepsilon(a) \right\}.$$

For an observer, the value of ε required to derive an optimal decision a^* is not available, so the observer can only consider and sum up each possible value of (random variable) ε weighted by its corresponding probabilities. This integrated optimal policy (for observers) is thus probabilistic and is called a conditional choice probability (CCP), denoted by:

$$\pi(a|x) \triangleq P(a \in \arg \max_a \{Q(x, a') + \varepsilon(a')\}).$$

In addition to *Assumptions 1 and 2*, we also assume ε follows an *i.i.d* multivariate Gumbel distribution, that is,

Assumption 3 (*i.i.d* Gumbel random perturbations). ε is *i.i.d* and follows multivariate Gumbel distribution, or:

$$\mu(d\varepsilon) = \prod_{a \in A} \exp\{-\varepsilon(a) + \gamma\} \exp\{\exp\{-\varepsilon(a) + \gamma\}\},$$

where $\gamma \approx 0.5772$ is the Euler's constant.

Then, the expected cumulative subjective value function over random perturbation ε has the closed-form expression:

$$V(x) \triangleq \int V(x, \varepsilon) d\mu(\varepsilon) = \gamma + \log \sum_a \exp Q(x, a).$$

Moreover, the CCP has the closed-form expression:¹²

$$\pi(a|x) = \frac{\exp Q(x, a)}{\sum_{a'} \exp Q(x, a')}.$$

Maximum Likelihood Estimation

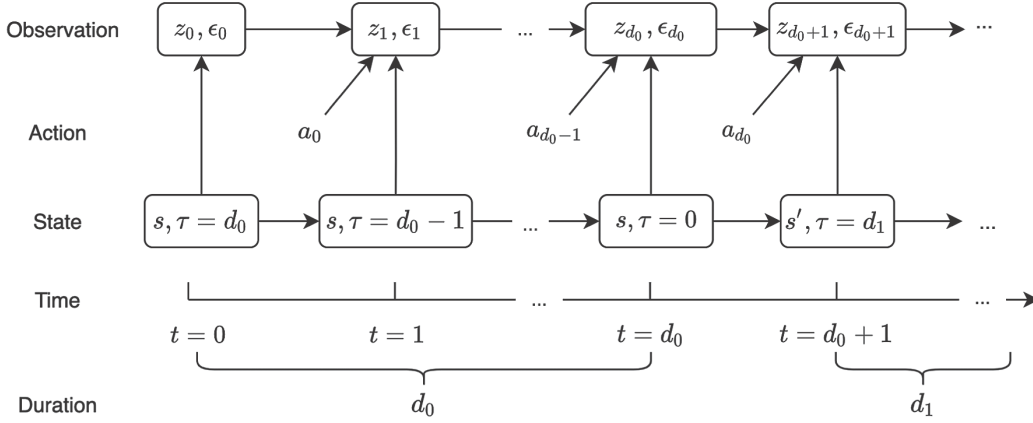
Our estimation goal is to identify the dynamics of mental fatigue $\{P_{\theta_D}(s', \tau', z|s, \tau, a) : s, s' \in \mathbb{S}, \tau, \tau' \in \mathbb{D}, z \in \mathbb{Z}, a \in \mathbb{A}\}$ and one-step subjective values $\{r_{\theta_R}(s, a) : s \in \mathbb{S}, a \in \mathbb{A}\}$ from data (i.e., ζ_T), whereby we parameterize fatigue dynamics by vector θ_D and subjective value by vector θ_R . Ideally, researchers should use data from experiments with human subjects. In this article, we will use simulated data to illustrate the estimation procedure developed in this section. It is well known that DDC models are in general not identifiable (i.e., it is impossible to uniquely recover model parameters from data). Our proposed model, in Equations 2 and 5, entails DDC models and hence is no exception. However, Chang et al. (2020) provided the following conditions to uniquely determine the model primitives.

¹⁰ The proofs of these existing theoretical results can be found in Chang et al. (2020).

¹¹ In practice, one stops at a finite T once some termination criterion is met.

¹² The relationship between the expected cumulative subjective value function and CCP have been discussed widely in past work (Chang et al., 2020; McFadden, 1980; Rust, 1994).

Figure 4
Structure of Proposed Model



Note. Arrow lines show directional statistical dependence. The hidden fatigue state s persists for a (random) sojourn time d_i , $i = 0, 1, \dots$. At each time, an observation z_t is obtained. The agent's subjective value of performing tasks with effort level a is perturbed by a random variable $\epsilon_t(a)$. The change in fatigue state s is affected by effort allocation choice a (at $\tau = 0$).

1. The dynamics of mental fatigue $\{P_{\theta_D}(s', \tau', z | s, \tau, a) : s, s' \in \mathbb{S}, \tau, \tau' \in \mathbb{D}, z \in \mathbb{Z}, a \in \mathbb{A}\}$ can be estimated from data if
 - a. Each data sequence contains two (or more) periods of data;
 - b. Both initial belief x_0 and the cardinality of fatigue state space S and longest duration in a fatigue state D (i.e., $|\mathbb{S}|, |\mathbb{D}|$) are known.¹³
2. One-step subjective values $\{r_{\theta_R}(s, a) : s \in \mathbb{S}, a \in \mathbb{A}\}$ can be identified from data if the following quantities are known:
 - a. The cardinality of state space \mathbb{S} and \mathbb{D} ;
 - b. Cumulative distribution function of random perturbation μ ;
 - c. Discount factor β ;¹⁴
 - d. one-step subjective values for a reference action a^0 , $\{r(s, a^0), \forall s \in \mathbb{S}\}$. That is, one-step subjective values are only identified up to a constant (i.e., only the differences in $r(s, a) - r(s, a^0)$ can be estimated, not the absolute values).

Under these conditions, maximum likelihood estimation can be used to find the value of θ_D, θ_R that jointly maximizes the likelihood function of given observable histories ζ_T s. Assume we have N independently generated histories in the form of data $= \{(x_0^n, a_t^n, z_{t+1}^n), 0 \leq t \leq T\}_{n=1}^N$, with each history containing $T + 1$ decision epochs. The log-likelihood function can be factorized by reduced-form observation probabilities and CCPs:

$$\begin{aligned} \log \mathcal{L}(\theta_D, \theta_R | \text{data}) &\triangleq \prod_{n=1}^N \prod_{t=0}^T \log \left[P_{\theta_D}(z_{t+1}^n | x_t^n, a_t^n) P_{\theta_D, \theta_R}(a_t^n | x_t^n) \right] \\ &= \sum_{n=1}^N \sum_{t=0}^T \left[\log \sigma_{\theta_D}(x_t^n, a_t^n, z_{t+1}^n) + \log \pi_{\theta_D, \theta_R}(a_t^n | x_t^n) \right]. \end{aligned}$$

The estimation can be split into two phases. In the first phase, we estimate θ_D by maximizing a partial log-likelihood function regarding reduced-form observations because it only contains θ_D . The maximum likelihood estimator is denoted by $\hat{\theta}_D$:

$$\hat{\theta}_D \in \arg \max_{\theta_D} \sum_{n=1}^N \sum_{t=0}^T \log \sigma_{\theta_D}(x_t^n, a_t^n, z_{t+1}^n).$$

In the second phase, we use $\hat{\theta}_D$ obtained from phase 1 as a given and maximize a partial likelihood function containing CCPs:

$$\hat{\theta}_R \in \arg \max_{\theta_R} \sum_{n=1}^N \sum_{t=0}^T \log \pi_{\hat{\theta}_D, \theta_R}(a_t^n | x_t^n),$$

wherein the maximum likelihood estimator of θ_R is denoted by $\hat{\theta}_R$.¹⁵ Figure 5 depicts our proposed modeling and estimation framework. Note that we further assume

$$P_{\theta_D}(s', \tau', z' | s, a, \tau) = P_{\theta_1}(s', \tau' | s, a, \tau) P_{\theta_2}(z' | s'),$$

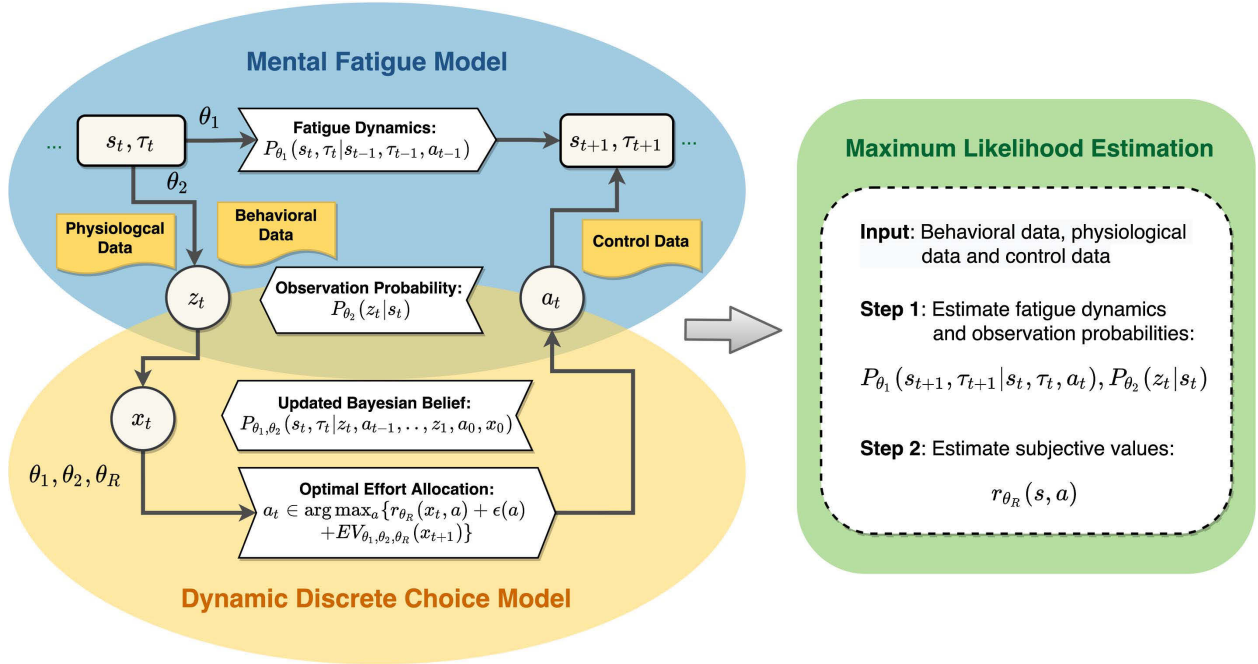
where $\theta_D = (\theta_1, \theta_2)$. Physiological and behavioral measures correlate with underlying fatigue states [although this correlation is far from perfect (Bijleveld, 2018)]. Thus, we model the observation as a conditional probability given current fatigue state $P(z | s)$, which is commonly called an observation probability.

¹³ Usually x_0 can be obtained from self-reports. Numerical examples also show that the influence of x_0 on estimation results is increasingly negligible as the number of decision epochs T increases (e.g., see Chang et al., 2020).

¹⁴ The values of $|\mathbb{S}|, |\mathbb{D}|, \beta$ and distribution μ can be treated as design parameters and a practitioner can test different options on these parameters for a best fit.

¹⁵ Theoretically, we should jointly maximize θ_D, θ_R for the optimal likelihood function. However, the joint estimation is computationally expensive. In practice, a partial estimation (first estimating θ_D and then estimating θ_R based on the estimator of θ_D)—such as our proposed procedure—is widely used and usually yields good performance (i.e., consistent, asymptotically normally distributed estimators; Cox, 1975; Rust, 1994).

Figure 5
Modeling of Mental Fatigue and Its Effects on Effort Allocation



Note. The Blue circle contains an abstraction of changes or transitions in mental fatigue states. The intertemporal dynamics of both the hidden fatigue state and its duration are parameterized by θ_1 . Noisy measurements (such as behavioral or physiological data) of fatigue states are denoted by observations and are stochastically dependent on the underlying hidden fatigue state. This dependent structure is captured by $P_{\theta_2}(z_t | s_t)$. Decisions refer to the agent's choice of difficulty level (i.e., effort allocation) for the upcoming block of the cognitive task, which manifests a subject's preference for the level of effort to be exerted and results in changes to both fatigue state and its duration. The Yellow circle depicts our proposed model for describing mental effort allocation. The model assumes that the agent exerts mental effort to maximize the expected cumulative subjective value (immediate value plus future expected value) of successful task performance, which is determined by fatigue dynamics, observation probabilities, and one-step subjective values (parameterized by θ_R). This subjective value is a function of the fatigue state at each decision epoch. Because fatigue states and their duration are hidden, the model assumes that the agent self-perceives his or her fatigue state and its duration, described by the belief x . The Green box displays the estimation procedures for the model parameters, including fatigue dynamics (θ_1), observation probabilities (θ_2), and one-step subjective values (θ_R). Specifically, we first do partial estimation regarding fatigue dynamics ($P_{\theta_1}(s, \tau | s, \tau, a)$) and observation probabilities ($P_{\theta_2}(z | s)$). Then, we use the estimators of fatigue dynamics and observation probabilities to estimate one-step subjective values ($r_{\theta_R}(s, a)$). See the online article for the color version of this figure.

More accurate measurements will be more likely to reflect the true fatigue state.

N-Back Task

In this section, we apply the proposed model and estimation procedure to a simulated n -back task. We begin by explaining important model components and expressing them in mathematical terms. Then we give concrete values to the model parameters for a simulated n -back task. A fully specified model allows us to simulate rational mental effort allocations and changes in fatigue, and to illustrate the proposed estimation procedure. We also discuss the estimation results and validate the estimated model by comparing its predictions to empirical results from the relevant literature on mental fatigue and effort expenditure. Last, we perform robustness analysis for cases wherein required prior knowledge for the model to function is unknown.

Fatigue and Mental Effort Allocation During an N-Back Task

Task setup

We consider a free-choice task environment composed of several blocks of the n -back task, which is a task commonly used to manipulate mental fatigue. We assumed that higher levels of task difficulty (i.e., larger n) lead to higher payoffs (e.g., monetary rewards) for successful performance but also carry higher psychological costs. For each block, the agent's choice of task difficulty level ($a_t = n$, $n = 1, 2, 3$, wherein $n = 1$ is the easiest version of the task) and fatigue-related measures (z_t) will be recorded, hence, $\zeta_{t+1} = \{z_t, a_t, z_{t+1}\}$.¹⁶

¹⁶ In an actual experiment, fatigue measures could include electroencephalography, electrocardiography, performance metrics, etc. Here, in a simulated environment, the fatigue measures were composed of random variables stochastically dependent on hidden fatigue states.

Model of Mental Fatigue Dynamics

A common way to induce mental fatigue is to have subjects perform cognitively demanding tasks for a prolonged period of time (Boksem et al., 2005, 2006; Hopstaken et al., 2015, 2016; Wylie et al., 2020). Studies have observed that performing more demanding tasks leads to greater mental fatigue compared to performing less demanding tasks (Blain et al., 2016; Wylie et al., 2017). To that end, task difficulty and time on task are factors worth investigating in the proposed model. We modeled changes in mental fatigue as a hidden semi-Markov process:

1. $s_t \in \mathbb{S} = \{1, 2, \dots, S\}$: all possible values for fatigue state at time t . A larger value of s implies a more fatigued state.
2. $\tau_t \in \mathbb{D} = \{1, 2, \dots, D\}$: time left in current state s_t at time t , the maximum of which is D . Ideally, D should be less than the duration of the experiment so that subjects become fatigued before the experiment ends.

Then we have

$$P(s_{t+1}, \tau_{t+1}, z_{t+1} | s_t, \tau_t, a_t), \quad (7)$$

to describe transitions between fatigue states at consecutive decision epochs.

The One-Step Subjective Values $r(s, a)$

The subjective value $r(s, a)$ considers the trade-off between rewards $b(s, a)$ and costs $c(s, a)$ for a single decision period:

$$r(s, a) = b(s, a) - c(s, a).$$

Note that by *Assumption 2*, random perturbations in subjective values ϵ are usually not explicitly stated, yet softmax optimal policy used in extant studies of effort allocation (Chong et al., 2017; Hartmann et al., 2013; Klein-Flügge et al., 2015; Prévost et al., 2010) were derived using certain assumptions for the distribution of ϵ (i.e., under the assumption that $\epsilon(a)$ is *i.i.d* across different actions and follows a Gumbel distribution, optimal action $a \in \arg \max_{a'} \{r(s, a') + \epsilon(a')\}$ is probabilistic with softmax form).

Cost Function $c(s, a)$. To model the psychological costs for performing an n -back task, we could use the cost functions defined by existing work on cognitive control (Chong et al., 2017; Hartmann et al., 2013; Klein-Flügge et al., 2015; Prévost et al., 2010), including:

$$\begin{aligned} \text{Linear :} & \quad c(s, a) = ksa; \\ \text{Hyperbolic :} & \quad c(s, a) = \frac{1}{1-ksa}; \\ \text{Parabolic :} & \quad c(s, a) = k(sa)^2; \\ \text{Exponential :} & \quad c(s, a) = e^{ksa}; \end{aligned}$$

where $k > 0$ is the parameter characterizing individual or group differences and a is the task difficulty (see Figure 6). In general, the

parameter a in the cost function $c(s, a)$ refers to effort rather than task difficulty. However, a positive correlation exists between task difficulty and the amount of effort to be exerted, particularly when task success is deemed likely (Brehm & Self, 1989; Fairclough & Ewing, 2017).¹⁷

Reward Function $b(s, a)$. According to motivation theories and extant studies of cognitive control, the expected rewards for mental effort expenditure play an important role in guiding human behavior. In addition, subjective beliefs may influence the perceived value of external rewards. Subjective beliefs that may influence perceived value include self-efficacy and success probability (or expectancy) (Atkinson, 1957; Bandura, 1977; Feather, 1959a, 1959b; Maier & Seligman, 1976; Rotter, 1966; Wabba & House, 1974). Despite differences in terminology, in general, all these terms relate to the likelihood of succeeding at a prospective task. For example, the motivation to pursue a task or goal may be quantified by weighing possible task incentives with their associated success likelihoods, that is, Classic expectancy \times Value considerations (Reinhardt & Wabba, 1975). Efficacy is crucial in the evaluation of subjective value, and thus influences effort allocation (Froemer et al., 2020).

For example, Chow et al. (2015) found that participants report reduced self-efficacy for task-related exertions after performing a mentally demanding task. Here, we define self-efficacy beliefs as the subjective evaluation of the probability of succeeding at a given task:

$$e(s, a) = P(\text{success} | s, a),$$

where $e(s, a)$ links both task difficulty a (effort level) and fatigue state s to success probability. In terms of the n -back task, self-efficacy describes how confident a subject is in being able to succeed (i.e., meet performance criteria). According to existing research and theory, we expected $e(s, a)$ to decrease as the task becomes more difficult or as the subject becomes more fatigued. Consequently, the reward function reflects external incentive values combined with self-efficacy beliefs:

$$b(s, a) = e(s, a) \text{incentive}(a; c_p),$$

where $\text{incentive}(a; c_p)$ is the monetary payoff subjects would receive upon meeting the (predetermined) performance criteria c_p for the task a . Performance criteria c_p are specified by the experiment design and will be omitted in the function $b(s, a)$ for brevity.

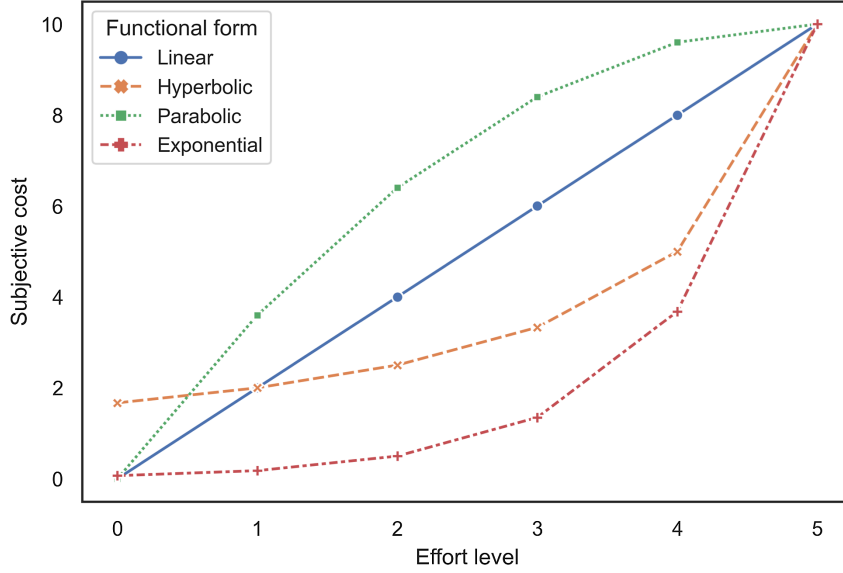
Expected Cumulative Subjective Value Function

The one-step subjective value $r(s, a)$ is the subjective value for a specific decision epoch. The optimal allocation of effort for a

¹⁷ In human subjects, task difficulty can be linked to effort with careful experimental design. For example, subjects can be informed that incentives will be earned only if certain performance standards are met. Hence, choosing a task that is beyond one's capability to achieve (i.e., the performance standard cannot be met) is irrational because the rewards are unobtainable. Such an experiment design helps guide subjects to choose task difficulty levels that match the effort they are willing to exert.

Figure 6

Commonly Used One-Step Subjective Costs, Increasing as Perceived Effort Increases



Note. Typical forms of subjective costs include linear, hyperbolic, parabolic, and exponential. See the online article for the color version of this figure.

forward-looking agent is to maximize the expected cumulative subjective values over horizon T :

$$\{a_0, a_1, \dots, a_T\} \in \arg \max_{a_0, \dots, a_T} E \left[\sum_{t=0}^T \beta^t (r(s_t, a_t) + \varepsilon(a_t)) | \zeta_0 \right],$$

where $\beta \in [0, 1)$ is the discount factor. The operator E denotes the expectation over all stochastic components: (a) all future actions yet to be selected; (b) all future random perturbations ε ; and (c) all future fatigue states and their duration times (s, τ) that transition stochastically.

Simulated N -Back Tasks With Changing Fatigue

We assume that the current fatigue state is either 1 ($s = 1$ for *not fatigued*) or 2 ($s = 2$ for *fatigued*). The temporal duration of staying in one of these fatigue states τ takes possible values from 0 to 4 (blocks of trials).¹⁸ We assume the following conditional independence of the dynamics of fatigue state and its duration:

$$\begin{aligned} P(s_{t+1}, \tau_{t+1}, z_{t+1} | s_t, a_t, \tau_t) \\ = P(s_{t+1} | s_t, a_t, \tau_t) P(\tau_{t+1} | s_{t+1}, a_t, \tau_t) P(z_{t+1} | s_{t+1}), \end{aligned}$$

and that:

1. Fatigue is more likely to occur when the agent chooses to perform (and then performs) harder tasks. Here we assume that choosing to perform more difficult tasks of $n \geq 2$ when $\tau = 0$ (about to get tired) will fatigue the agent with probability 1, that is, $P(s_{t+1} = 2 | s_t, \tau_t = 0, a_t \neq 1) = 1, \forall s_t$. Choosing the easiest task $a_t = 1$ will help the agent to recover, so $P(s_{t+1} = 1 | s_t, \tau_t = 0, a_t = 1) = 1, \forall s_t$. Note that when $\tau > 0$, $s_{t+1} = s_t$ regardless of choice a_t .¹⁹

2. Performing harder tasks will increase the probability of transitioning to the fatigued state, and the agent will also need more time to recover from that fatigued state. For example, $P(\tau_{t+1} = 4 | s_{t+1} = 2, \tau_t = 0, a_t = 3) \geq P(\tau_{t+1} = 4 | s_{t+1} = 2, \tau_t = 0, a_t = 2)$ implies that for the next state, the probability of staying in a fatigued state for five blocks (the longest possible duration) is greater when the current choice was to perform the 3-back task, compared to the 2-back task. Note that when $\tau > 0$, $\tau_{t+1} = \tau_t - 1$ regardless of choice a_t .

For (one-step) subjective values, we use a linear cost function (with respect to effort level) and a weighted reward function (weighted by self-efficacy beliefs), which is parameterized by $\theta_3 = \{\theta_{3,1}, \theta_{3,2}\}$ ²⁰:

$$r(s, a) \triangleq b(s, a) - c(s, a) = \frac{\theta_{3,1}}{sa} \text{incentive}(a) - \theta_{3,2}sa,$$

where we assume self-efficacy has the form $e(s, a) = \frac{\theta_{3,1}}{sa}$, and where $\theta_{3,1}, \theta_{3,2}$ are positive real numbers to be estimated as part of the model parameters.

To simulate an n -back task environment, we had to predetermine the value of the model parameters. There are no prior studies incorporating both a free-choice paradigm and mental fatigue due to performing the n -back task. Nevertheless, Hopstaken et al. (2015) provided a referable benchmark. They studied the effects of mental fatigue on task performance during a prolonged series of n -back

¹⁸ Note that in Hopstaken et al. (2015), the subjects were fatigued within 5–6 blocks of trials. Here, we similarly assumed that fatigue can be induced within five blocks of trials, indicating that the maximal number of time one can stay in a not fatigued state is 5 (from Block 0 to Block 4).

¹⁹ We will revisit those assumptions in the Limitations section to consider cases wherein the assumptions may not hold.

²⁰ The same procedure can be applied for nonlinear cost functions.

tasks. The authors observed changes in task performance (measured in d -prime), physiological states (P3 amplitude, pupil diameter), and self-reports of fatigue. We thus adapted data from their experiment to define our simulation environment. Note that free choice over n -back task difficulty was not included in Hopstaken et al., which could have been used by our model to identify subjective values from participants' behaviors.

1. We parameterize the transition between fatigue states $P(s'|s, a, \tau = 0)$ in a deterministic fashion because only two states are possible; leaving one state necessarily leads to entering the other. This parameter is treated as known and thus did not need to be estimated.²¹ Specifically, we have:
 - a. $P(s' = 2|s, \tau = 0, a) = 0, a = 1, \forall s$;
 - b. $P(s' = 2|s, \tau = 0, a) = 1, a = 2, 3, \forall s$;
 - c. $P(s' = 1|s, \tau = 0, a) = 1, a = 1, \forall s$;
 - d. $P(s' = 1|s, \tau = 0, a) = 0, a = 2, 3, \forall s$.
2. The sojourn time distribution is described by θ_1 , or $P_{\theta_1}(\tau|s', a, \tau = 0)$, where $\theta_1 = \{\theta_{1,1}, \theta_{1,2}, \theta_{1,3}\}$ (see definitions below). We assume that sojourn times follow a binomial distribution. Note that the binomial distribution takes two parameters, the second of which is defined as the maximum number of blocks a state can persist, which is fixed to 4.²² This follows from the experiment by Hopstaken et al. (2015), in which subjects exhibited signs of fatigue within five consecutive blocks of n -back task performance.
 - a. $P(\tau|s' = 1, a = 1, \tau = 0) \sim B(\theta_{1,1} = 0.1, 4)$;
 - b. $P(\tau|s' = 2, a = 2, \tau = 0) \sim B(\theta_{1,2} = 0.4, 4)$;
 - c. $P(\tau|s' = 2, a = 3, \tau = 0) \sim B(\theta_{1,3} = 0.7, 4)$.

We designed average sojourn time in the *fatigued* state as longer for $a = 3$ than for $a = 2$.²³

3. The observation probability—which characterizes the relationship between a hidden fatigue state s and its observable noisy measurement z —is captured by θ_2 , or $P_{\theta_2}(z|s)$. We used the performance data reported in d -prime Hopstaken et al. (2015) as the observation z . Hopstaken et al. observed that average performance across subjects consistently declined.²⁴ Performance (d -prime) levels on the relevant n -back tasks observed by Hopstaken et al. fluctuated between 2.4 and 3.2. To make our model outputs and predictions comparable to their results, we borrowed their mean values for the highest performance level (d -prime ≈ 3.12 among not fatigued participants) and lowest performance level (d -prime ≈ 2.54 among fatigued participants). Furthermore, we added variance to the performance variable to account for the fact that individual performance rarely matches the group-level mean. We constrained performance (d -prime) within a range [2.5, 3.5] because most of the findings reported by Hopstaken et al. fell within that range. We remapped the observation z from the discretized d -prime in Table 1 and defined the observation probability $P(z|s)$ in Table 2.

Table 1

We Mapped Performance d -Prime Into Integer Space

Observation	d -prime \in [2.5, 2.75)	d -prime \in [2.75, 3)	d -prime \in [3, 3.25)	d -prime \in [3.25, 3.5)
z	1	2	3	4

Note. For example, when d -prime is in the range [2.5, 2.75), corresponding z takes value 1.

4. We describe the one-step subjective value function by $\theta_R = \theta_3 = \{\theta_{3,1}, \theta_{3,2}\} = \{2.5, 0.5\}$.
5. We specified the external incentives (rewards) earned for each block of n -back as: One unit of reward for the 1-back task, 3 units of reward for the 2-back task, and 6 units of reward for the 3-back task.

The model parameter (to be estimated) of $\theta = \{\theta_1, \theta_2, \theta_3\}$ is summarized in Table 3. The discount factor $\beta = 0.9$, subjective value perturbation ε is assumed to be *i.i.d* and Gumbel distributed (to ensure the model can be estimated). Now the proposed fatigue model is fully specified.

Model Estimation, Validation, and Results Analysis

We simulated 1,000 sample paths with each path containing 40 consecutive blocks of trials (or 40 decision epochs) for the given value of θ . Recorded history includes choices (or actions) of the task difficulty levels and simulated task performance.

Estimation results are displayed in Table 3. Our estimation method was able to identify both fatigue dynamics and the one-step subjective value function with high precision. The maximum element-wise difference between the ground truth of the parameters and their estimates was 0.01 for θ_1 , θ_2 and θ_3 , which differed by no more than 4% relative to the ground truth.

In addition, the estimated model was able to validate classic findings from existing empirical studies. Figure 7 compares the changes in fatigue due to prolonged n -back task performance reported by Hopstaken et al. (2015) to the predictions from our estimated model.²⁵ As time on task increased, Hopstaken et al.

²¹ For the number of states greater than two, transition probabilities cannot be treated as known; they will be estimated as model primitives. See an example in Robustness Analysis section.

²² Note that in Hopstaken et al. (2015), blocks were coded from 1 (e.g., the first block of trials was labeled as Block 1). In this article, we code from Block 0 (i.e., the first block of each sample path is Block 0). This decision aligns with the fact that binomial distribution can take 0 as a value. As a result, the second parameter of binomial distributions in the article is 4 (rather than 5), which indicates a maximum sojourn time of five blocks (from Block 0 to Block 4).

²³ In the parameterization of sojourn time distributions, we considered only cases ($s' = 1, a = 1$), ($s' = 2, a = 2$), ($s' = 2, a = 3$), because the aforementioned fatigue state transitions rule out other combinations. For example if $a = 1$, then $s' \neq 2$. Similarly, $s' \neq 1$ if $a = 2, 3$.

²⁴ In their experiment, 3-back performance was consistently poor and showed a slight increase as time on task proceeded. We ignored their 3-back task findings in defining the range of performance values because, according to Hopstaken et al., 2015 the increase was likely induced by learning effects, which are not a focus of the current work.

²⁵ Whereas our approach considers a free choice paradigm, Hopstaken et al. (2015) did not. Nonetheless, our model was able to generate predictions comparable to the data observed in Hopstaken et al.

Table 2
Parameterization of Observation Probability by θ_2

s	$z = 1$	$z = 2$	$z = 3$
$s = 1$	$\theta_{2,1,1} = 0.01$	$\theta_{2,1,2} = 0.10$	$\theta_{2,1,3} = 0.79$
$s = 2$	$\theta_{2,2,1} = 0.69$	$\theta_{2,2,2} = 0.28$	$\theta_{2,2,3} = 0.02$

Note. Note that $P(z = 4|s) = 1 - P(z = 1, 2, 3|s)$, which does not require estimation.

observed that participants' subjective feelings of fatigue increased in a concave manner (faster increases at the beginning and slower increases toward the end of the task). Our model revealed a similar increase in fatigue levels and captured the concave shape as well (see the declining slope of the blue solid line at Blocks 5 and 6). Note that our model expresses subjective fatigue in terms of probabilities, that is, $P(s = 2|\zeta)$, whereas Hopstaken et al. gathered subjective reports of fatigue.

In Figure 8 we compared 2-back task performance observed by Hopstaken et al. (2015) to our estimated model outputs. Performance (d -prime) declined as time on task increased in Hopstaken et al.; see orange dashed line. Our model used performance as a noisy indicator of underlying hidden fatigue states. As time on task increased, the model predicted declines in task performance. The predicted performance decline was faster for the first several blocks, after which time the declines slowed down (flattened); see blue solid line.²⁶

Furthermore, Hopstaken et al. (2015) found declines in task engagement with increased time on task, but apparently reengagement when external incentives were promised for good performance (Figure 9, orange dashed line). This pattern suggested that incentives alter the subjective value of mental effort expenditure and thus make expending effort more attractive. Such incentive-induced change can also be predicted by our model. To mimic the way, Hopstaken et al. incentivized subjects for good performance, we increased base incentives for the 2- and 3-back tasks by 80%. The model predicted that agents become more prone to expend effort for potential rewards even though fatigued; see Figure 9 (blue solid line). In our model, task engagement was defined by the probability of selecting more demanding tasks (2-back and 3-back), that is, $\pi(a = 2|x) + \pi(a = 3|x)$. Moreover, as subjective values kept increasing, our model predicted that the marginal gains in task engagement diminish (see Figure 10). This finding is aligned with evidence that the marginal utility of rewards decreases once the rewards exceed a certain level (Grahek et al., 2020; Kahneman & Tversky, 1979; Morgenstern & Von Neumann, 1953; Pine et al., 2009).

Robustness Analysis

Our model assumes that some of the model primitives are given, namely, the discount factor β and the number of fatigue states S . However, those assumptions may be overly restrictive relative to real life. For example, what is the number of hidden fatigue states that can best describe a subject's effort allocation? Or what is the intertemporal discount factor a subject uses when making decisions? In this section, we will analyze our model's estimation performance in two concrete cases wherein the cardinality of state space (S) or the intertemporal discount factor is (β) unknown.

Unknown Value of β

The discount factor is involved in the computation of forward-looking subjective value functions. Variations in the discount factor may lead to (potentially large) differences in the value functions and thus affect the estimation of subjective value parameters and individual differences in decision-making. The identification of the discount factor (from data) in dynamic discrete choice models is a theoretical challenge and an active area of research (e.g., Abbring & Daljord, 2020). It has been shown that the discount factor is not identifiable from decision choice data alone without further restrictions (Magnac & Thesmar, 2002; Rust, 1994). Usually, the identification of a discount factor requires additional context-dependent information (e.g., Bajari et al., 2016; Norets & Tang, 2014).

To examine the effects due to unknown discount factor, we considered a case wherein the model settings (fatigue dynamics, subjective value parameters, etc.) remain the same as the simulated n -back task in the previous section, except that the discount factor β is unknown. To this end, we generated data using the discount factor $\beta = 0.7$. Then, we performed the proposed parameter estimation procedure assuming different true values of $\beta \in [0, 1)$. To compare the model estimates and their true values, we denoted the estimates for certain β by $\hat{\theta}_3(\beta) = \{\hat{\theta}_{3,0}(\beta), \hat{\theta}_{3,1}(\beta)\}$ and the ground truth by $\theta_3 = \{\theta_{3,0}, \theta_{3,1}\}$. (Note that phase 1 estimation of dynamics parameters does not involve β .) Then, the maximum element-wise percentage difference is defined as:

$$D(\theta_3, \hat{\theta}_3(\beta); \beta) = 100 \times \max_{i \in \{1, 2\}} \left\{ \frac{|\theta_{3,i} - \hat{\theta}_{3,i}(\beta)|}{\theta_{3,i}} \right\}$$

Figure 11 shows how the estimated log-likelihood and the difference (%) between estimated subjective value function parameters $\hat{\theta}_3$ and their true values θ_3 change with respect to β . The maximum log-likelihood and the minimum of $D(\theta_3, \hat{\theta}_3; \beta)$ were achieved simultaneously at the true β value ($\beta = 0.7$), as would be expected. From the figure, see also that when the assumed β is close to the true β ($= 0.7$), the estimation result is acceptable in terms of both performance (log-likelihood, blue solid line) and the precision of the estimates (orange dashed line). For example, as the assumed value of the discount factor ranged between 0.6 and 0.8, there was no significant drop in model performance and the estimated percentage difference (compared to the true values) was under 20%, attesting the robustness of our approach. The simulation results shown in Figure 11 imply that β values can be fine-tuned by numerical methods (such as uniformly searching values between 0 and 1), which represents a viable way to apply our model when the true discount factor is unknown.

In addition to a numerical searching method to find a proper value of the discount factor, another route for alleviating the reduced

²⁶ The task setup from Hopstaken et al. (2015) was not exactly the same as the task we attempted to model; thus, we did not observe exactly the same "dip" in task performance of Block 4 found by Hopstaken et al., 2015 Figure 3, 2-back task. Specifically, Figure 3 in Hopstaken et al. (2015) was plotted by averaging data across all subjects, whereas our model is specified for an individual. In addition, their experiment paradigm did not involve free task selection. Thus, their data cannot be directly incorporated into our model. Rather, we used their data as a reference point and modified accordingly to fulfill our needs for establishing simulations. For example, we used their task performance (d -prime) range to determine the range of observations in our model.

Table 3
Model Primitive Parameters and Their Estimates

Parameter	θ_1			θ_2						θ_3	
	$\theta_{1,1}$	$\theta_{1,2}$	$\theta_{1,3}$	$\theta_{2,1,1}$	$\theta_{2,1,2}$	$\theta_{2,1,3}$	$\theta_{2,2,1}$	$\theta_{2,2,2}$	$\theta_{2,2,3}$	$\theta_{3,1}$	$\theta_{3,2}$
True value	0.1	0.4	0.7	0.01	0.10	0.79	0.69	0.28	0.02	2.5	0.5
Estimated value	0.10	0.39	0.70	0.01	0.10	0.79	0.68	0.29	0.02	2.51	0.50
Average difference		0.003				0.003				0.005	

identifiability caused by an imprecise discount factor would be to gather additional (separate) evidence for a participant’s (or a sample’s average) discount factor. For example, performance on a classic delay discounting task could suggest a beta value on which to anchor the relevant model estimation parameter for understanding the impact of mental fatigue on effort allocation decisions. Please also refer to the Limitations section, wherein we address issues pertaining to the temporal discounting of subjective value.

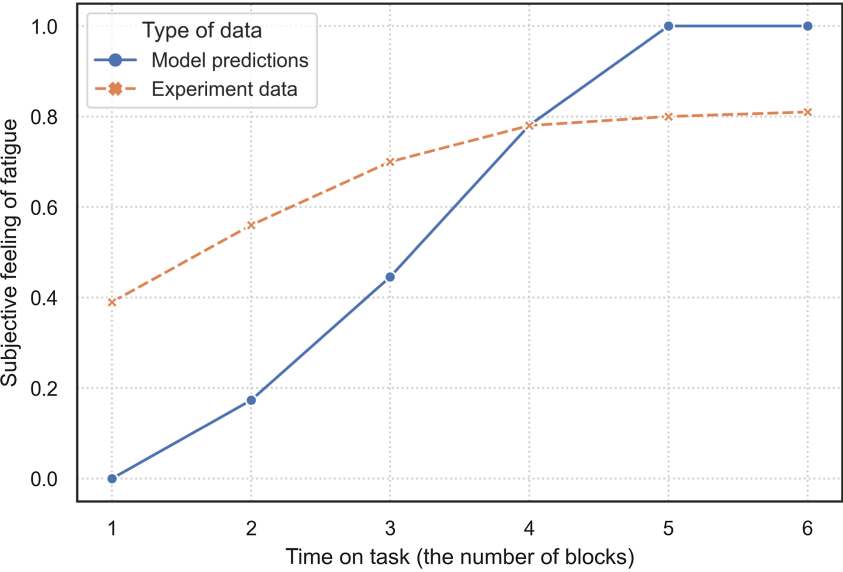
Unknown Cardinality of Fatigue State Space

Mental fatigue is a multifaceted construct that may not be easily categorized or quantified. However, one can still find a reasonable set of fatigue categories or states that can be used to explain an agent’s effort expenditure in terms of modeling. In principle, the choice of the number of fatigue states should be kept parsimonious, because more states may lead to more parameters required to specify the model, and the increase in model parameters demands larger amounts of data to ensure the precision of model estimates. Akaike information criterion (AIC) and

Bayesian information criterion (BIC) have been used to evaluate the quality of a model, which take into account the trade-off between model complexity and model fit. Usually, the model with lower AIC/BIC values is preferred. For illustration purpose, we considered the following case, wherein simulated data was generated by a 3-state model. Assuming a modeler does not know the cardinality of fatigue state space, the fact that fatigue state space cardinality is discrete allows the modeler to search over a subset of possible numbers of states to find an appropriate (well-fitting) number. We followed this approach by parameterizing our model using 2-state, 3-state, and 4-state settings, respectively. The model settings can be found in Appendices A–C.

The estimation results (log-likelihood) are on display in Figure 12 and in Table 4. As can be seen in Figure 12, increasing the number of fatigue states from 2 to 3 resulted in a marked improvement in the likelihood, whereas increasing from 3 to 4 states led to no more gain. Table 4 shows the AIC and BIC values for the different models, wherein the 3-state model yielded both the smallest AIC and BIC. To summarize, for a data set generated by a 3-state model, the best estimates were achieved by a 3-state setting, which clearly

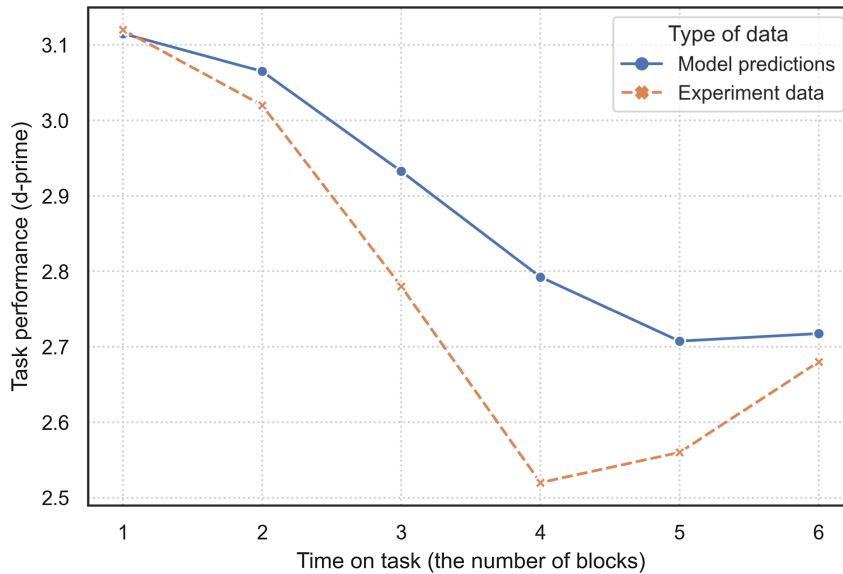
Figure 7
Fatigue Increased With Time on Task, as Reported in a Past Study and as Predicted by Our Proposed Model



Note. Whereas our approach considers a free-choice paradigm, Hopstaken et al. (2015) did not. Nonetheless, our model was able to generate predictions comparable to the data observed in Hopstaken et al.. The orange dashed line represents the findings reported by Hopstaken et al. (2015) whereas the blue solid line shows the simulated results from our proposed model. See the online article for the color version of this figure.

Figure 8

Task Performance (Measured in d -Prime) Decreased as Time on Task Increased as Observed in Prior Work [adapted From (Hopstaken et al., 2015, Figure 3, 2-Back Task) and Our Model Predictions



Note. Whereas our approach considers a free-choice paradigm, Hopstaken et al. (2015) did not. Nonetheless, our model was able to generate predictions comparable to the data observed in Hopstaken et al. The orange dashed line represents the findings from Hopstaken et al. (2015) whereas the blue solid line shows our model predictions. See the online article for the color version of this figure.

outperformed a 2-state setting. Although a practitioner may not know the “true” number of underlying hidden states, he/she can experiment with different settings and select the setting with the smallest AIC/BIC value to achieve closest-to-true-model performance.

Discussions

We modeled the allocation of effort under mental fatigue as a DDC model with hidden and semi-Markov states. This model was based on reward/cost trade-offs, treated effort expenditure as capacity-limited (rather than resource-limited), and assumed that people behave rationally by seeking to maximize cumulative subjective values across the whole decision cycle. We posited that subjective value is influenced by mental fatigue and provided a way to calculate cumulative subjective values using some common assumptions. Moreover, we developed a maximum likelihood estimation procedure to identify model parameters—one-step subjective values and mental fatigue dynamics—from an agent’s decisions and behavioral or physiological data. We tested the proposed estimation method with a simulated free choice n -back task composed of several blocks of trials. Simulated agents had to choose a task difficulty level for the upcoming block while taking into account the effects of mental fatigue. We fed simulation data to the estimation procedure and found high estimation precision in identifying model parameters. Further, we successfully reproduced some canonical findings from mental fatigue and decision-making literature using predictions from the estimated model, including decreased task performance with increased time on task as well as increased feeling of fatigue, task reengagement with additional

performance incentives, and reduced marginal effects of further effort exertion with added incentives.

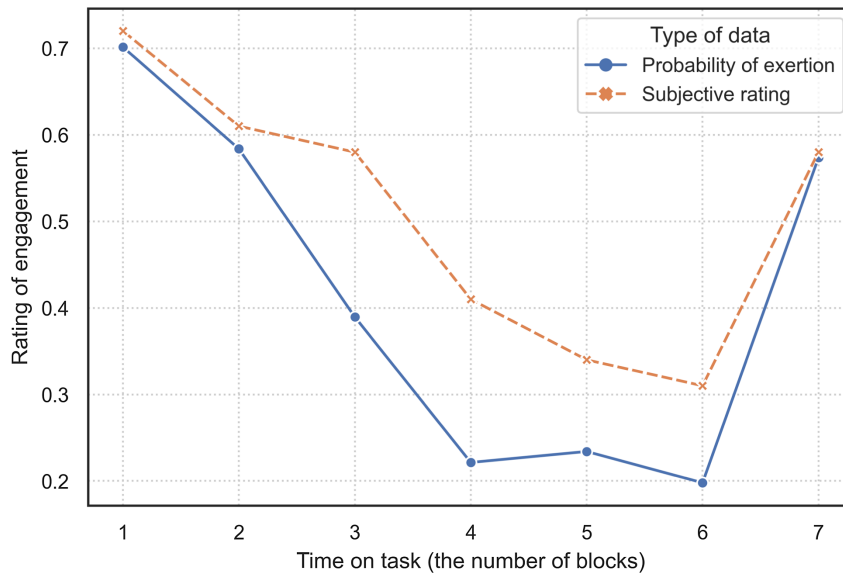
Below, we consider related work and summarize our approach’s unique theoretical implications. We then envision some prospective applications that may benefit from our approach and also address some limitations of the proposed model.

Related Work

Some prior studies have addressed cognitive control and mental fatigue based on a reward/cost trade-off account. Those studies provided both empirical and theoretical insights into mechanisms that may account for mental effort exertion. Here we compare the current work with prior studies to reveal the unique contribution of our approach.

EVC theory was proposed to explicate the principles of cognitive control allocation (Shenhav et al., 2013, 2017). We sought to build on the EVC model by incorporating the influence of mental fatigue on effort allocation decisions. In the EVC model, the decision to be made regarding cognitive control is twofold: the identity of control (i.e., what kind of cognitive control to engage) and the intensity of control. A rational agent needs to optimally adjust these control signals to maximize the expected value. A computational study illustrated the EVC model in the context of a simulated flanker task based on a one-step control allocation problem (Musslick et al., 2015). The environment state was composed of sensory stimuli, which indicated the direction of the arrows in the flanker task. The control allocation decision of the agent was assumed to be based on the inferred state

Figure 9
Task Engagement From Hopstaken et al. (2015) and From Our Model



Note. Whereas our approach considers a free-choice paradigm, Hopstaken et al. (2015) did not. Nonetheless, our model was able to generate predictions comparable to the data observed in Hopstaken et al. *Orange dashed line* depicts the changes in self-reported subjective feeling of task engagement adapted from Figure 2 in Hopstaken et al. (2015). From Blocks 1–6, engagement decreased while mental fatigue accumulated. And the willingness for further exertion reverted in Block 7 due to an external incentive guaranteed by the good performance. *Blue solid line* exhibits the model simulated changes in the willingness to allocate effort as a function of time on task. We had simulated agents perform 2-back task for consecutive seven blocks of trials. From Blocks 1–6, the model predicted a decrease in willingness to exert effort, which was then reversed in Block 7 by introducing an external incentive for good performance [similar to the settings used in Hopstaken et al. (2015)]. Specifically, when the extra monetary rewards equaled 80% of current monetary rewards for both 2-back and 3-back tasks, the predicted willingness to expend effort (i.e., the probability of choosing either the 2-back or 3-back task for the next block) reverted to (near) prefatigue levels, resulting in a probability increase from 0.2 (in Block 6) to 0.58 (in Block 7). See the online article for the color version of this figure.

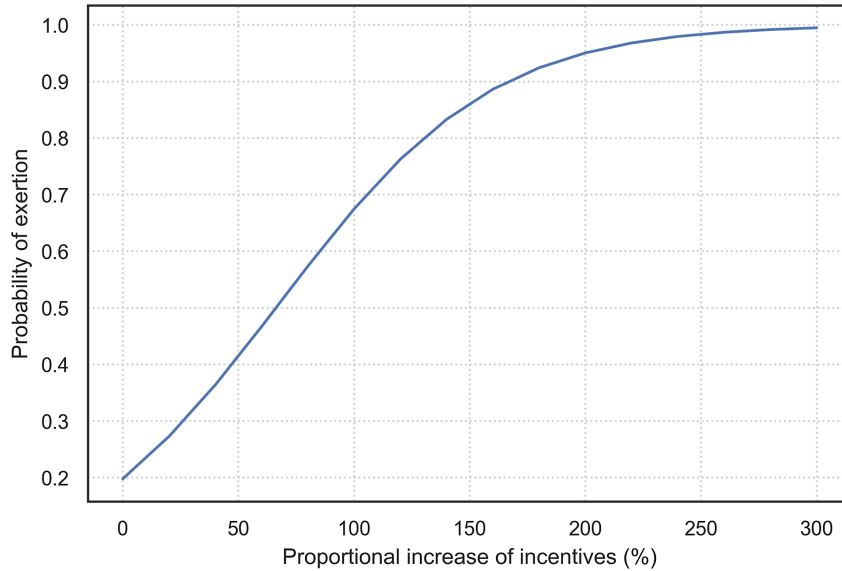
(the subjective representation of task stimuli),²⁷ and then the environment would emit an outcome (e.g., correct or incorrect response) determined by both the true state (the actual direction of the target arrow) and the subject's control decision (i.e., response). The update of automatic bias (e.g., the tendency to be influenced by flanking arrows rather than target arrows) was based on the difference between emitted outcomes (from the true state of the environment) and anticipated outcomes (from the inferred state of task stimuli), weighted by a learning rate. The model used by Musslick et al. involved a relatively short (trial to trial) time scale. In the current work, we focused on longer term subjective states (mental fatigue) that may influence effort allocation decisions. Further, differences in mathematical formulations used in the prior work and the current work can be found. First, we considered the calculation of subjective values for an infinite horizon, which avoided the (possibly) overly simplified (one step) formulation of subjective value. Second, we took a Bayesian brain perspective, modeling self-perception of mental fatigue as based upon the entire history of relevant observations and choices, rather than updating an inferred state (e.g., automatic bias) based on current perceptions and the difference between anticipated and actual outcomes as in Musslick et al.

EVC theory has also been applied to the study of the influences of subjective states on cognitive control allocation, such as affect. More specifically, Grahek et al. (2020) studied the influence of affect on cognitive control by applying the EVC framework to simulated Stroop tasks. They parameterized the costs and rewards to build subjective values and based the analysis on the simulation results. In addition to the different research targets (i.e., affect vs. fatigue), distinctions can be found between that work and the current model. We modeled mental fatigue as a hidden state that plays an important role in building subjective values and hidden state dynamics. Grahek et al. *indirectly* studied the influence of affect. For example, to study the influence of integral affect (i.e., a type of affect that arises from task performance and related incentives), Grahek et al. increased rewards for successful performance and compared the simulation results (e.g., task performance, willingness to exert cognitive control) to a model setting with smaller task

²⁷ In a flanker task, subjects indicate which direction a target arrow points, with the target arrow surrounded by other nontarget arrows (which may aim in a different direction than the target). The nontarget (flanking) arrows may distract attention—making subjective representation of task stimuli noisy and inaccurate—and may elicit an incorrect response to the target.

Figure 10

The Marginal Effects of the Willingness to Expend Effort (i.e., the Probability of Choosing the 2-Back or 3-Back Task) as a Function of Increased Task Rewards



Note. The starting point of the curve is the probability of effort expenditure in simulated agents (built by the estimated model) who have performed the 2-back task for six blocks of trials. We manipulated the external incentives for successful performance for both 2-back and 3-back tasks in terms of proportionality of the increase (compared to current received rewards). We found that task engagement was increased by external rewards and grew rapidly as the rewards increased by between 0% and 150%. However, the marginal effects of task engagement diminished when external incentives exceeded a certain level (e.g., more than 200%) as depicted by the reduced slope of the curve. See the online article for the color version of this figure.

rewards. Although that approach permitted the researchers to perform simulations and compare the differences between two model settings, it also raises questions such as *how to determine the appropriate magnitude of task rewards?* and *how to address individual differences?* (because different people may react differently to the different task reward values). By contrast, we provided a way to address those questions by estimating model parameters (subjective values and mental fatigue dynamics) from experimental data.

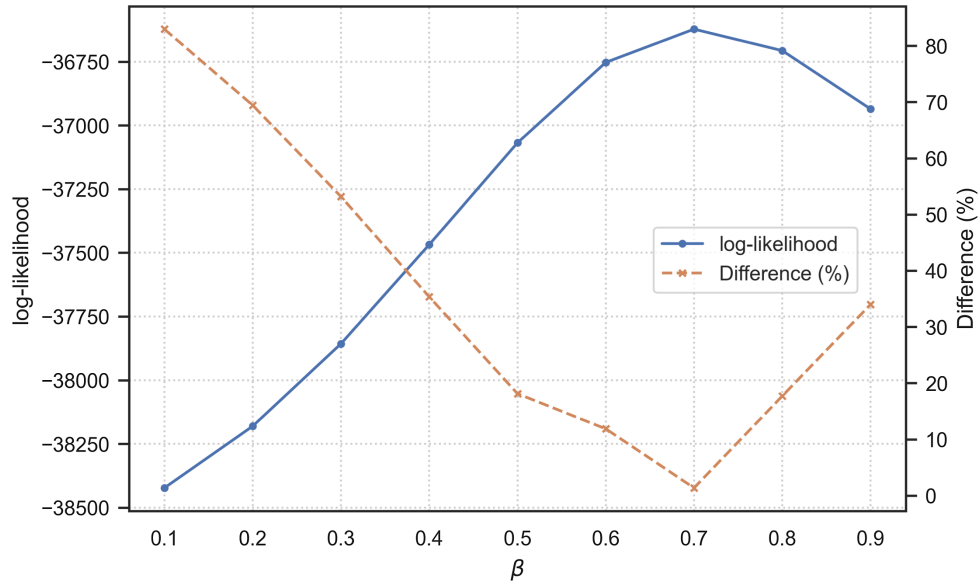
Estimation of the cost function using the EVC model was discussed in Musslick et al. (2018). They considered a one-step forward-looking formulation of EVC. In other words, cognitive control (e.g., identity and intensity) specified at time t results in some outcome that carries value at time $t + 1$. The optimal control policy should satisfy the first-order condition by taking partial derivatives of EVC with respect to control. The derivatives contained two parts: one part was the derivative of the cost function (related to current step values), and the other part involved the derivatives of task accuracy (related to future values). Hence, by first-order optimality condition, the cost function could be estimated via the derivatives of task accuracy, and those derivatives can be obtained from specific experimental designs (such as by manipulating task rewards and observing the corresponding changes in task accuracy to estimate the desired derivatives). In that work, a one-step forward-looking formulation of EVC and the implicit assumption that the values of outcomes did not depend on the current decision reduced the complexity of finding first-order conditions. In the current work,

the choice at time t affects future value by altering the agent's fatigue state; future value is the expected cumulative value for all decisions (at time $t + 1$, $t + 2$, ...) yet to come (rather than looking only one-period ahead). Our proposed estimation procedure proved able to deal with this relatively more complicated, nested form of subjective value function (hence, a recursive optimal effort allocation policy).

Agrawal et al. (2020) proposed a model based on utility maximization to explain why people become (mentally) fatigued and bored. They proposed that fatigue and boredom bias human behavior toward more rewarding activities. In the case of fatigue, people may quit online cognitive operations in favor of offline computations (such as mental simulation) to improve the likelihood of gaining future rewards. In the case of boredom, people may quit the task at hand to search for more rewarding tasks. By specifying the agent's subjective preference (subjective utility) under different circumstances [different state-action pairs (s, a)], the model proposed by Agrawal et al. can be used to explain why and when fatigue and boredom occur. However, specifying an agent's subjective utilities is not a trivial undertaking. Our work aims directly at this issue: We estimate the agent's subjective values by their choices.²⁸ In other words, we do not assume the reasons behind task disengagement; we only assume that mental fatigue biases

²⁸ The relationship between subjective preference and choices has been considered in the theory of revealed preferences, which states that humans' choices are the best indicator of their preferences, see Samuelson (1938).

Figure 11
Estimation Performance Under Different Values of β



Note. Blue solid line shows the model performance (in terms of log-likelihood) with respect to different β . Orange dashed line shows the maximum element-wise percentage difference between estimates under different β and ground truth. See the online article for the color version of this figure.

behavior by altering subjective values, and we can fit our model to data to quantify such fatigue-induced biases. In addition, our model treats mental fatigue states as partially observable, allowing for noise associated with fatigue measures. Our proposed estimation procedure can still identify underlying hidden fatigue dynamics with high precision based only on noisy (imprecise) measures.

Theoretical Implications

The power of learning from data endows the capability and potential of the proposed approach. We begin discussion of the model's implications by stressing (a) our model is an instance of a more general estimation methodology (structural estimation of POMDPs), which needs empirical data to function and (b) the model is not limited to a specific form of subjective value function or fatigue dynamics; it invites researchers to test and compare their theories related to the constructions of subjective values as well as the dynamics of mental fatigue. In what follows, we provide some examples of how our model can contribute to the study of mental fatigue.

Distinguish Rewards From Costs

In the current model, the cost of mental effort $c(s, a)$ and the perceived reward of successful task performance $b(s, a)$ are modeled jointly in the one-step subjective value function as $r(s, a) = c(s, a) - b(s, a)$. The current method will identify $r(s, a)$; however, one could parameterize the cost function $c(s, a)$ and the perceived reward function $b(s, a)$ using evidence from empirical studies. The model could then estimate the effect of mental fatigue on the cost and perceived reward functions separately, which could help to understand the extent to which effort allocation decisions under mental fatigue stem mainly from changes in the perceived costs of

expending effort, or the perceived rewards of expending effort, or some combination of both factors. Such an undertaking could shed light on the extent to which the perceived value of rewards changes under mental fatigue (Kelley et al., 2019).

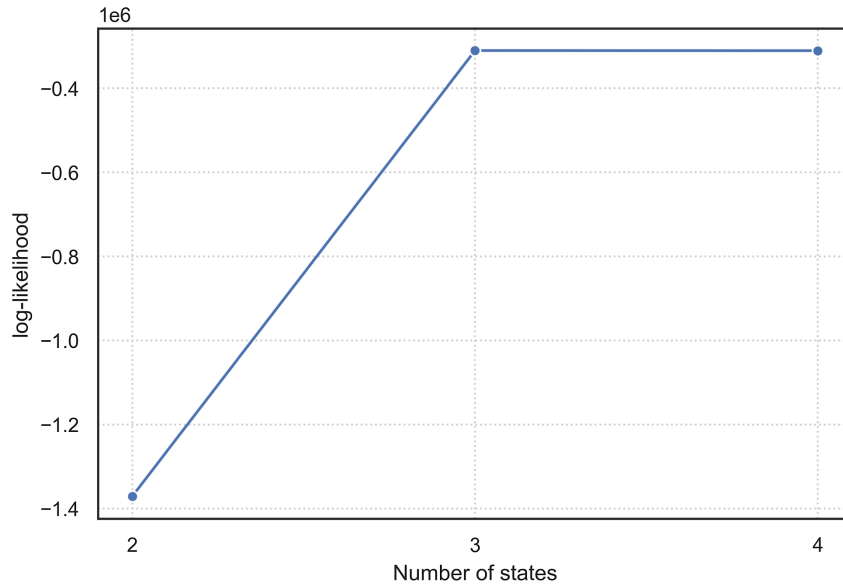
Are Subjective Reports of Mental Fatigue Objectively Accurate?

The measurement of mental fatigue usually includes both objective (e.g., task performance, physiological records) and subjective (e.g., self-reports) indicators. Unfortunately, the two indicators do not always align. Debates over the relative merits of the two indicators are common (Holtzer et al., 2017; Sandry et al., 2014; Völker et al., 2016). The modeling and estimation approach described here provides a possible way to address this issue: by fitting the model with objective measures (e.g., performance or physiological changes), one could compare the estimated dynamics of mental fatigue with subjective self-reports, which may help to specify the relationship between subjective and objective measurements of mental fatigue. It then will be interesting to examine for whom and under what circumstances subjective and objective indicators of mental fatigue cohere, and when they fail to cohere. We propose that research and theory on mental fatigue would benefit by identifying break points in coherence (i.e., circumstances wherein the two indicators diverge) and by identifying demographic, personality trait, or brain morphological factors that relate to coherence (or lack thereof) between subjective and objective indicators of mental fatigue.

Future Applications

Our modeling and estimation procedure to identify hidden fatigue dynamics and subjective values may prove useful for both theoretical and empirical studies of mental fatigue and effort allocation.

Figure 12
Model Estimation Performance Under Different Numbers of Fatigue States



Note. Log-likelihood of simulated data set was computed from estimated models including 2-, 3-, or 4-state parameterizations. See the online article for the color version of this figure.

A Perspective to Deal With Circularity

A crucial question raised by Botvinick and Braver (2015) regarded the mechanisms behind reward/cost trade-offs in explaining human behavior: Behavioral preferences are explained by subjective values, but subjective values, in turn, are usually inferred from behavioral preferences. The situation grows more complicated when mental effort carries costs: More effort leads to both more rewards and more costs. Hence, a clear pathway for obtaining subjective values is needed to break the circular reasoning identified above and thus ensure that “the assertion that choices maximize values becomes potentially falsifiable and thus truly scientific” (Padoa-Schioppa, 2011). Botvinick and Braver further claimed that one way to overcome the circularity is to leverage findings from neuroscience to consolidate the construction of subjective values. Relevant candidates include physiological indicators such as electroencephalography, pupil diameter, and neural activations related to effort allocation, subjective evaluation, and mental fatigue dynamics (Bartra et al., 2013; Boksem et al., 2005; Hopstaken et al., 2016;

Padmala & Pessoa, 2010; Schmidt et al., 2012). One study used functional magnetic resonance imaging (fMRI) to test the hypothesis that incentives can enhance task performance by improving task information coding and maintenance (Etzel et al., 2016). Efficiency in decoding task information was significantly higher under incentive compared to no-incentive trials. Although neural activity correlated qualitatively with incentive evaluation, one question worth probing is: How much improvement can be achieved by introducing incentives? Or similarly, how much difference does an incentive make in terms of subjective values? In the study of mental fatigue, the need for such quantitative analysis exists as well. For instance, as discussed previously, Hopstaken et al. (2015) found that the promise of rewards reversed the negative effects of mental fatigue and psychophysiological indicators associated with task disengagement. But how much reversal of such negative effects can be accomplished by offering additional rewards remained unexplained.

Our modeling and estimation procedure paves a new way for the rigorous and objective quantification of subjective values based on the behavioral, neural, physiological, and decision-making outputs of the decision-maker. This quantification can be achieved by incorporating neuroscientific findings into the proposed framework. For example, in a task-switching paradigm (Etzel et al., 2016), subjective value functions could be constructed via extracted fMRI patterns (in this example, fMRI activity corresponds to cue and early delay periods). If, for the simplest case, we could parameterize subjective values by a linear function of fMRI signals, then by estimating the corresponding parameters in the linear function, we may be able to build a quantitative relationship between choice preferences and their neural correlates via subjective values. Such parameterization enables one to build subjective value functions without undertaking the risk of circularity, and to simplify the estimation of subjective value functions by focusing on a few design

Table 4
AIC, BIC of Different Models for Simulated Data

Model selection criteria	<i>S</i> = 2	<i>S</i> = 3	<i>S</i> = 4
	No. parameters = 11	No. parameters = 20	No. parameters = 32
AIC	2,740,992	620,613	621,281
BIC	2,741,101	620,812	621,601

Note. AIC = Akaike information criterion; BIC = Bayesian information criterion. We calculated AIC and BIC values for estimated models with different numbers of hidden states. Note that for the different models, the number of parameters required to describe the model is different.

parameters. In addition to the parameterization of subjective value functions, our framework features a customizable module for researchers to plug in the fatigue measurements they prefer (for the modeling of observation probability), which is a (conceptually) simple computation to implement. Yet it can be supercharged by neuroscientific advancements: More reliable fatigue measures lead to less noisy observations, which in turn enhance precision in the estimation of fatigue dynamics and subjective values. For example, recent years have witnessed the emergence of nonintrusive physiological measurements (such as smart wearables) of people's health status, which provide abundant observable data that could facilitate fatigue estimation. Researchers have already begun to realize the potential of smartphone or smart wearable data in assessment and monitoring of psychological states (Kodithuwakku Arachchige et al., 2021). Specifically, vital signs such as heart rate and heart rate variability, which may be tracked via smartwatches or bands, are useful in predicting mental fatigue (Luo et al., 2020). Further, eye blink data can be collected using smart glasses and may also contribute to detecting mental fatigue (Horiuchi et al., 2018).

Our model may take the advantage of aforementioned technology breakthroughs and help researchers answer quantitative questions such as “*how much* performance gain can be achieved for a given incentive?” Lastly, our model is capable of describing within or across individuals variations in the mapping from neural correlates to subjective values, which will be discussed in the next section. Simply put, with a sufficient amount of data, our model can be fitted separately for different individuals as well as groups of people, which may facilitate the study of individual-level and group-level perceptions of subjective value.

Individual Differences

Individual differences in personality traits (Braem et al., 2013; Jimura et al., 2010), age (Ennis et al., 2013), and clinical disorders (Marx et al., 2013) are likely to impact subjective values (Westbrook & Braver, 2015). Differences in subjective values may predict differential willingness to allocate mental effort in daily life (Westbrook et al., 2013). One study of cognitive control using a two-step decision-making model found significant individual differences in the tendency to incorporate goal-related contextual information in decision-making (Otto et al., 2015). Such individual differences likely play a role in effort allocation decisions as well. Age is another likely moderator of mental effort allocation; it is possible that aging alters the perceived costs of expending mental effort (Westbrook et al., 2013). In the study of mental fatigue, individual perceptions play an important role as well. Evidence has suggested that older people may be influenced by mental fatigue in a different manner compared to young people, and older people may fatigue more quickly and suffer greater performance decrements compared to young people (Arnau et al., 2017).

Our model provides a way to study the influence of such individual differences: The estimation procedure can identify the one-step subjective values (and thus, reward function and cost function parameters) and fatigue dynamics for each individual or group. By comparing estimation results across individuals or groups, one could attempt to quantify the influence of the individual difference of interest. Furthermore, a counterfactual analysis (i.e., comparing what actually happened against what would have happened under different conditions) could be used to predict individual differences in reward/cost sensitivity (Braem et al., 2013; Engelmann et al., 2009). Our model is able to easily simulate changes in incentives and observe

the consequences of such changes, which could then be tested (verified) in experiments with human subjects.

Perceptions of effort and mental fatigue may be important also for understanding some clinical disorders. The perceived costs of expending effort may relate to specific clinical conditions. People with depression or schizophrenia may experience more effort-related costs compared to others (Westbrook et al., 2013). Fatigue descriptions in clinical practice are various and have different emphases in different contexts: some focus on both physical fatigue and mental fatigue (e.g., chronic fatigue syndrome, see Pattyn et al. (2008) and Shen et al. (2006)) whereas others may stress mental fatigue in particular (e.g., insomnia, see Pattyn et al. (2008)); some can be alleviated by sleeping (Mairesse & Neu, 2016) whereas others are more acute and may not be improved by rest (Hossain et al., 2003). Furthermore, the intensity of perceived fatigue may differ across individuals with underlying clinical conditions. As mental fatigue is a multifaceted construct, it is unlikely that one particular measure can precisely and comprehensively capture a person's level of mental fatigue (Pattyn et al., 2008).

Our model introduces a novel way for clinical studies to approach effort expenditure and mental fatigue. In addition to having the capacity to incorporate relevant physiological and self-report measures of mental fatigue, our model infers subjective values from behavioral choices. In clinical conditions that implicate effort deficits, analysis of subjective values has proved successful in predicting symptoms such as substance abuse, obesity, and problem gambling (Westbrook & Braver, 2015). In the evaluation of mental fatigue in clinical populations, a “transactional model” of fatigue has been proposed to measure levels of mental fatigue based on the patient's experience (Pattyn et al., 2008). The proposed model suggests a practical implementation of this idea by fitting our mental fatigue model to each individual patient, resulting in a unique clinical profile capturing personal effort exertion willingness and changes in willingness due to mental fatigue. The potential exists for this approach to help tailor treatments to individuals for mitigating effort- and fatigue-related decrements.

Controllability and Task Performance

It has been proposed that work performance depends on both how demanding a task is and the degree of controllability the person has over his or her activities (Hockey, 2011; Hockey & Earle, 2006). Empirical results have suggested that people may sustain relatively high levels of work performance without reporting fatigue if they are allowed to arrange the work at their discretion (Frese, 1989; Karasek, 1979; Wall et al., 1996). In a more recent study, Hockey and Earle built a simulated office work environment wherein the workload was characterized by both time pressure and controllability (of task schedule). Under high workload, only the low controllability group exhibited negative effects on work performance and fatigue.

Our model provides a way to study individual preferences for Demand \times Control interactions. An adapted free-choice paradigm permits an agent to allocate effort by distributing workloads to specific periods of time. Assuming that a decision-maker behaves rationally and thus, given a defined workload and a specific time span, optimally allocates effort based on subjective values, researchers would be able to quantify how subjective values change as the agent grows more fatigued and how the agent maintains a balance

between work progress and the negative effects of mental fatigue. This balance is important because it marks the optimal effort allocation strategy (for the agent) and can be explained by a compensatory control model (Hockey & Earle, 2006) such that the discrepancy between work goals and effort exertion are reconciled. The interpretation from our model would be that the agent adjusts his or her fatigue state by altering the allocation of effort, thereby keeping fatigue levels in an acceptable range with economically optimal trade-offs between costs and rewards.

Dissociation of Choice Dimensions

Dissociability is an important concept in decision-making and neuroeconomics. Both the delay of rewards and the uncertainty of receiving rewards conspire to reduce subjective values. Generally speaking, more delayed rewards and less certain rewards are valued less than their sooner or more certain counterparts. How can one be certain whether delay or uncertainty is responsible for an observed discounting of rewards? Further, how might we distinguish discounts induced by effort from discounts induced by delay or uncertainty? If these discount factors cannot be dissociated, then effort-based decision-making may be just another form of delay- or uncertainty-based decision-making (Westbrook & Braver, 2015). Researchers have designed different experimental paradigms to dissociate the costs of effort from the costs from delay or uncertainty. In the current work, we have introduced fatigue as another potential factor that may serve to discount subjective values. Moreover, although not addressed in the current model, factors other than fatigue could influence the subjective evaluation of task values, including such as negative moods or boredom (Chiew & Braver, 2011; Pattyn et al., 2008; Pessoa, 2008; van Steenbergen et al., 2012).

As the number of the variables that influence subjective values increases, it becomes harder to dissociate their effects by pure experimental design techniques. Our model opens the door to dissociate effects of different variables via mathematical optimization of the likelihood of data: The variables of interest can be included in the reward/cost functions and estimated. For example, agents may manifest both “overload” (limited cognitive capacity) and “underload” problems (often described as boredom; Hancock & Desmond, 2001). It has been argued that although both types of problems represent a form of fatigue (sometimes described as active fatigue and passive fatigue, respectively), they have different mechanisms. It may be informative to examine the discounting mechanisms associated with overloads and underloads—by parameterizing subjective value as a function of both passive and active fatigue—to understand how they influence subjective values that may explain the agent’s effort allocation behaviors. We intend our model to be used in the future research to untangle the processes underlying the outcomes of mental fatigue.

Limitations

Simplifying Assumptions Revisited

In the Simulated N -back Tasks with Changing Fatigue, we illustrated our method via a simulated n -back task under the assumption that choosing to perform the harder task hastens fatigue (which is supported by prior studies, e.g., Blain et al., 2016). This assumption was linked to the idea that the agent earns the reward associated with the task only if

their performance surpasses a threshold presumed to entail some degree of mental effort (particularly when the more difficult task was chosen). If the agent expends little or no effort on the difficult task, then it is likely that no reward is earned; the simpler task would have been the more rational choice. Nevertheless, human agents do not always make rational decisions. Another assumption we made was that choosing to perform easier tasks restores mental fatigue, though this assumption also may not hold in all circumstances. For example, tasks that do not entail much mental effort, instead of restoration, may be prone to elicit boredom. We cannot rule out the possibility that agents will grow bored, rather than or in addition to mentally fatigued, when performing n -back tasks. In fact, a previous study of reward sensitivity and boredom found it hard to fully decouple fatigue and boredom (Milyavskaya et al., 2019).

To be sure, our model does not present an exhaustive analysis for each possible scenario. The aforementioned assumptions were adopted only for illustrative purposes and are based on prior evidence consistent with the assumptions. Such assumptions influence the ways we parameterized fatigue dynamics and subjective value functions, but the same estimation procedure would still apply to obtain those estimates. Furthermore, one could also fit the data under different assumptions to see which assumptions best explain observed behaviors, thereby providing a way to test the validity of the assumptions. This consideration represents another example of how researchers can use our model and estimation methods in future work to advance research and theory on mental fatigue.

Concerns Arising From the Rationality Assumption

Our proposed model may serve many future studies of mental effort and mental fatigue, though the model carries the inherent liabilities of the reward/cost trade-off framework. One concern is the assumption that people behave rationally (Rust, 1994). In many scenarios, people may act based on local contexts by comparing available options rather than evaluating all possible options. As a result, the relative values of different options may change with the immediate task environment (Delgado et al., 2011; Tversky, 1969). In our proposed model, the agent is assumed to evaluate the subjective values for each option and select the option with the highest subjective value (i.e., act independent of bias induced by local contexts). Another potential issue is that people may have difficulty evaluating the subjective values of states that are not currently occupied (Dayan, 2012; Loewenstein & O’Donoghue, 2004). For example, a person may underestimate the difficulty of being on a diet when he or she is currently not hungry. If relative incapability to predict nonoccupied states is a general principle in mental effort allocation, then our proposed model may not be able to provide satisfying explanations or predictions for people’s behaviors, because we assumed that human agents are rational and can objectively evaluate the values of future outcomes rooted from current state and choice selection.

Affect and Effort Allocation

Our current work focused on the “cognitive” side of mental fatigue and its effects, leaving the “affective” side relatively untouched. Emotions (or affects, used interchangeably from here on) are ubiquitous and play an important role in human decision-making, whether or not people are aware of their influence.

However, how emotions influence decision-making is still under debate. Different theories have been proposed. For example, some researchers have posited a valence-based approach [to what extent an emotion is experienced as positive/negative matters (Han et al., 2007; Loewenstein & Lerner, 2003)] whereas others have proposed an arousal-based approach [the intensity of emotions matters (Kaufman, 1999; Sohn et al., 2015)]. Some theorists believe a two-dimensional description of emotion is far from enough (Mellers et al., 1998) and have proposed an appraisal-tendency framework, which posits that emotion is a multidimensional construct and carries motivational factors that bias decision-making depending on both qualitative properties (which type of emotion) and quantitative properties (the intensity of such emotion). We notice that while being attributed to different reasons/mechanisms, different theories of emotion share a common feature: emotions carry value that can influence decision-making. Lazarus and Lazarus (1991) further argued that different emotions are related to different “themes” in assessing the value of events.

Following the assumption that emotions carry value, our model may prove useful for studying the influence of emotions on decision-making. The relevant state space in the model can incorporate both fatigue and affective states. As a result, subjective values would be a function of both fatigue and affective states; the temporal dynamics would describe the joint transition of those two states across time. Implementing such a model is not a trivial undertaking and was not the main focus of our present work, but the role of affect in shaping subjective value represents a novel route for future research to build on the model we have proposed. Another promising avenue for future work may be to consider compensatory mechanisms that help to sustain good performance under mental fatigue. A similar compensatory mechanism has been proposed to understand high performance in the context of cognitive and emotional stressors (e.g., Pavlidis et al., 2016), and it seems reasonable to assume that similar mechanisms exist to combat mental fatigue. These ideas represent possible applications of our model in the study of joint effects of mental fatigue and affect on decision-making. Because our approach is an instance of the structural estimation of POMDPs, one would need to deliberately design model parameters and experiment settings to unveil the power of the model (because it is meaningless to train a highly mis-specified model).

If, as discussed above, affect carries value, then some behaviors considered “irrational” by outsiders (not the decision-maker) may actually be rational to the decision-maker him/herself, in the sense that they contribute to subjective value. Our model is capable of dealing with hidden states, and thus may help rationalize decision-making behaviors by quantifying the values of such hidden values.

Modeling of Temporal Delay Discounting

Delay discounting, which is related to but potentially distinct from effort discounting, has long been of interest to researchers in the study, value-based decision-making. Delay discounting accounts for the phenomena whereby a distal reward (to be obtained in the future) is less valuable compared to the same amount of reward that can be obtained right now. A common finding, which is referred to as the amount effect, is that larger subjective values will be discounted less than smaller values (Green et al., 2013; Thaler, 1981). Individual differences have also been observed in delay

discounting rates: People who perceive higher costs tend to more heavily discount future values.

In the current work, we assumed that subjective values are discounted by β , which implies the proportionality of delay discounting does not depend on the amount of the subjective value. This assumption may not be appropriate for modeling future subjective values (i.e., it does not adequately account for the amount effect). And it seems to be reasonable to relate such amount effects to delay discounting mechanisms. Furthermore, recent studies have found that temporal delays may discount subjective value hyperbolically (Frederick et al., 2002; Green & Myerson, 2004; Laibson, 1997). In comparison, our model assumes an exponential discounting mechanism, suggesting that the discounting proportion of subjective values is determined by a discount factor β to the power of the duration of temporal delay $t(\beta^t)$. A worthwhile future research direction would be to incorporate different delay discounting mechanisms into the proposed model and compare which mechanism fits the data better.

Conclusion

We modeled the effects of mental fatigue on effort allocations based upon a reward/cost trade-off account. We posited that mental fatigue influences effort allocation decisions by altering subjective values. We provided an estimation method to identify mental fatigue dynamics and subjective values from experimental data. Our proposed model successfully verified mental fatigue-related phenomena such as performance drops, task disengagement, and reengagement with additional incentives. Further, our work offers the potential for advancing understanding of human effort allocation under mental fatigue, including estimation of individual differences, differentiating clinical conditions associated with compromised fatigue manifestations, investigations of neural underpinnings corresponding to mental fatigue and its influences on subjective value representations, and dissociability studies to isolate the effects of mental fatigue from multiple other determinants of subjective value. Collectively, our work, rooted in extant decision-making theories and prior studies of mental fatigue, unleashes a promising future of quantitative study of mental fatigue by harnessing the advantages of both empirical research and mathematical methods.

References

- Abbeel, P., & Ng, A. Y. (2004). Apprenticeship learning via inverse reinforcement learning. In R. Greiner & D. Schuurmans (Eds.), *Proceedings of the twenty-first international conference on machine learning*. Association for Computing Machinery. <https://doi.org/10.1145/1015330.1015430>
- Abbring, J. H., & Daljord, Ø. (2020). Identifying the discount factor in dynamic discrete choice models. *Quantitative Economics*, 11(2), 471–501. <https://doi.org/10.3982/QE1352>
- Agrawal, M., Mattar, M. G., Cohen, J. D., & Daw, N. D. (2020). *The temporal dynamics of opportunity costs: A normative account of cognitive fatigue and boredom*. BioRxiv.
- Aguirregabiria, V., & Mira, P. (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics*, 156(1), 38–67. <https://doi.org/10.1016/j.jeconom.2009.09.007>
- Arnau, S., Möckel, T., Rinkenauer, G., & Wascher, E. (2017). The interconnection of mental fatigue and aging: An eeg study. *International Journal of Psychophysiology*, 117, 17–25. <https://doi.org/10.1016/j.ijpsycho.2017.04.003>
- Arora, S., & Doshi, P. (2018). *A survey of inverse reinforcement learning: Challenges, methods and progress*. ArXiv:1806.06877.

- Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, 64(6), 359–372. <https://doi.org/10.1037/h0043445>
- Bajari, P., Chu, C. S., Nekipelov, D., & Park, M. (2016). Identification and semiparametric estimation of a finite horizon dynamic discrete choice model with a terminating action. *Quantitative Marketing and Economics*, 14(4), 271–323. <https://doi.org/10.1007/s11129-016-9176-3>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of bold fmri experiments examining neural correlates of subjective value. *Neuroimage*, 76, 412–427. <https://doi.org/10.1016/j.neuroimage.2013.02.063>
- Baum, L. E., & Petrie, T. (1966). Statistical inference for probabilistic functions of finite state markov chains. *The Annals of Mathematical Statistics*, 37(6), 1554–1563. <https://doi.org/10.1214/aoms/1177699147>
- Baum, L. E., Petrie, T., Soules, G., & Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains. *The Annals of Mathematical Statistics*, 41(1), 164–171. <https://doi.org/10.1214/aoms/1177697196>
- Bellman, R. (1966). Dynamic programming. *Science*, 153(3731), 34–37. <https://doi.org/10.1126/science.153.3731.34>
- Bijleveld, E. (2018). The feeling of effort during mental activity. *Consciousness and Cognition*, 63, 218–227. <https://doi.org/10.1016/j.concog.2018.05.013>
- Blain, B., Hollard, G., & Pessiglione, M. (2016). Neural mechanisms underlying the impact of daylong cognitive work on economic decisions. *Proceedings of the National Academy of Sciences*, 113(25), 6967–6972. <https://doi.org/10.1073/pnas.1520527113>
- Boksem, M. A., Meijman, T. F., & Lorist, M. M. (2005). Effects of mental fatigue on attention: An erp study. *Cognitive Brain Research*, 25(1), 107–116. <https://doi.org/10.1016/j.cogbrainres.2005.04.011>
- Boksem, M. A., Meijman, T. F., & Lorist, M. M. (2006). Mental fatigue, motivation and action monitoring. *Biological Psychology*, 72(2), 123–132. <https://doi.org/10.1016/j.biopsycho.2005.08.007>
- Botvinick, M., & Braver, T. (2015). Motivation and cognitive control: From behavior to neural mechanism. *Annual Review of Psychology*, 66, 83–113. <https://doi.org/10.1146/annurev-psych-010814-015044>
- Braem, S., Duthoo, W., & Notebaert, W. (2013). Punishment sensitivity predicts the impact of punishment on cognitive control. *PLOS ONE*, 8(9), Article e74106. <https://doi.org/10.1371/journal.pone.0074106>
- Brehm, J. W., & Self, E. A. (1989). The intensity of motivation. *Annual Review of Psychology*, 40(1), 109–131. <https://doi.org/10.1146/annurev.ps.40.020189.000545>
- Cassandra, A. R. (1994). *Optimal policies for partially observable markov decision processes* (Technical Report). Brown University.
- Chang, Y., Garcia, A., & Wang, Z. (2020). *Dynamic discrete choice estimation with partially observable states and hidden dynamics*. <https://arxiv.org/abs/2008.00500>
- Chiew, K. S., & Braver, T. S. (2011). Positive affect versus reward: Emotional and motivational influences on cognitive control. *Frontiers in Psychology*, 2, Article 279. <https://doi.org/10.3389/fpsyg.2011.00279>
- Choi, J., & Kim, K.-E. (2011). Inverse reinforcement learning in partially observable environments. *Journal of Machine Learning Research*, 12, 691–730. <https://dl.acm.org/doi/10.5555/1953048.2021028>
- Chong, T. T.-J., Apps, M., Giehl, K., Sillence, A., Grima, L. L., & Husain, M. (2017). Neurocomputational mechanisms underlying subjective valuation of effort costs. *PLoS Biology*, 15(2), Article e1002598. <https://doi.org/10.1371/journal.pbio.1002598>
- Chow, J. T., Hui, C. M., & Lau, S. (2015). A depleted mind feels inefficient: Ego-depletion reduces self-efficacy to exert further self-control. *European Journal of Social Psychology*, 45(6), 754–768. <https://doi.org/10.1002/ejsp.2120>
- Cox, D. R. (1975). Partial likelihood. *Biometrika*, 62(2), 269–276. <https://doi.org/10.1093/biomet/62.2.269>
- Dayan, P. (2012). How to set the switches on this thing. *Current Opinion in Neurobiology*, 22(6), 1068–1074. <https://doi.org/10.1016/j.conb.2012.05.011>
- Dayan, P., & Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, & Behavioral Neuroscience*, 8(4), 429–453. <https://doi.org/10.3758/CABN.8.4.429>
- de Jong, M., Bonvanie, A. M., Jolij, J., & Lorist, M. M. (2020). Dynamics in typewriting performance reflect mental fatigue during real-life office work. *PLOS ONE*, 15(10), Article e0239984. <https://doi.org/10.1371/journal.pone.0239984>
- Delgado, M. R., Phelps, E. A., & Robbins, T. W. (2011). *Decision making, affect, and learning: Attention and performance XXIII*. Oxford University Press.
- Diermeier, D., Keane, M., & Merlo, A. (2005). A political economy model of congressional careers. *American Economic Review*, 95(1), 347–373. <https://doi.org/10.1257/0002828053828464>
- Dittner, A. J., Wessely, S. C., & Brown, R. G. (2004). The assessment of fatigue: A practical guide for clinicians and researchers. *Journal of Psychosomatic Research*, 56(2), 157–170. [https://doi.org/10.1016/S0022-3999\(03\)00371-4](https://doi.org/10.1016/S0022-3999(03)00371-4)
- D’Mello, S., & Graesser, A. (2010). Modeling cognitive-affective dynamics with hidden markov models. *Proceedings of the Annual Meeting of the Cognitive Science Society*. CU Experts. https://experts.colorado.edu/display/pubid_228757#overview
- Doroshenko, L., Konyshov, V., & Selishchev, S. (2007). Classification of human sleep stages based on eeg processing using hidden markov models. *BioMedical Engineering*, 41(1), 25–28. <https://doi.org/10.1007/s10527-007-0006-5>
- Dvijotham, K., & Todorov, E. (2010). Inverse optimal control with linearly-solvable mdps. In J. Fürnkranz & T. Joachims (Eds.), *Proceedings of the 27th international conference on machine learning* (pp. 335–342). Omnipress.
- Eddy, S. R. (2004). What is a hidden markov model? *Nature Biotechnology*, 22(10), 1315–1316. <https://doi.org/10.1038/nbt1004-1315>
- Engelmann, J. B., Damaraju, E., Padmala, S., & Pessoa, L. (2009). Combined effects of attention and motivation on visual task performance: Transient and sustained motivational effects. *Frontiers in Human Neuroscience*, 3, Article 4. <https://doi.org/10.3389/neuro.09.004.2009>
- Engle, R. W. (2010). Role of working-memory capacity in cognitive control. *Current Anthropology*, 51(Suppl. 1), S17–S26. <https://doi.org/10.1086/650572>
- Ennis, G. E., Hess, T. M., & Smith, B. T. (2013). The impact of age and motivation on cognitive effort: Implications for cognitive engagement in older adulthood. *Psychology and Aging*, 28(2), 495–504. <https://doi.org/10.1037/a0031255>
- Etzel, J. A., Cole, M. W., Zacks, J. M., Kay, K. N., & Braver, T. S. (2016). Reward motivation enhances task coding in frontoparietal cortex. *Cerebral Cortex*, 26(4), 1647–1659. <https://doi.org/10.1093/cercor/bhu327>
- Fairclough, S. H., & Ewing, K. (2017). The effect of task demand and incentive on neurophysiological and cardiovascular markers of effort. *International Journal of Psychophysiology*, 119, 58–66. <https://doi.org/10.1016/j.ijpsycho.2017.01.007>
- Fairclough, S. H., & Houston, K. (2004). A metabolic measure of mental effort. *Biological Psychology*, 66(2), 177–190. <https://doi.org/10.1016/j.biopsycho.2003.10.001>
- Feather, N. T. (1959a). Subjective probability and decision under uncertainty. *Psychological Review*, 66(3), 150–164. <https://doi.org/10.1037/h0045692>
- Feather, N. T. (1959b). Success probability and choice behavior. *Journal of Experimental Psychology*, 58(4), 257–266. <https://doi.org/10.1037/h0046066>
- Finley, A. J., Tang, D., & Schmeichel, B. J. (2019). Sweet nothings: No effects of self-control exertion on blood glucose levels. *Social Psychology*, 50(5–6), 322–331. <https://doi.org/10.1027/1864-9335/a000376>
- Finn, C., Levine, S., & Abbeel, P. (2016). Guided cost learning: Deep inverse optimal control via policy optimization. In M. F. Balcan & K. Q. Weinberger (Eds.), *Proceedings of the 33rd international conference on machine learning* (pp. 49–58). Journal of Machine Learning Research.

- Frederick, S., Loewenstein, G., & O'donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, 40(2), 351–401. <https://doi.org/10.1257/jel.40.2.351>
- Frese, M. (1989). Theoretical models of control and health. In S. L. Sauter, J. J. Hurrell, & C. L. Cooper (Eds.), *Job control and worker health* (pp. 107–128). John Wiley & Sons.
- Friston, K. (2012). The history of the future of the bayesian brain. *Neuro Image*, 62(2), 1230–1233. <https://doi.org/10.1016/j.neuroimage.2011.10.004>
- Froemer, R., Lin, H., Wolf, C. K. D., Inzlicht, M., & Shenhav, A. (2020). *When effort matters: Expectations of reward and efficacy guide cognitive control allocation*. BioRxiv.
- Fu, R., Wang, H., & Zhao, W. (2016). Dynamic driver fatigue detection using hidden markov model in real driving condition. *Expert Systems with Applications*, 63, 397–411. <https://doi.org/10.1016/j.eswa.2016.06.042>
- Gailliot, M. T., & Baumeister, R. F. (2007). The physiology of willpower: Linking blood glucose to self-control. *Personality and Social Psychology Review*, 11(4), 303–327. <https://doi.org/10.1177/1088868307303030>
- Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D. M., Brewer, L. E., & Schmeichel, B. J. (2007). Self-control relies on glucose as a limited energy source: Willpower is more than a metaphor. *Journal of Personality and Social Psychology*, 92(2), 325–336. <https://doi.org/10.1037/0022-3514.92.2.325>
- Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, 349(6245), 273–278. <https://doi.org/10.1126/science.aac6076>
- Giacomantonio, M., Jordan, J., & Fennis, B. (2019). Intense self-regulatory effort increases need for conservation and reduces attractiveness of energy-requiring rewards. *Social Psychology*, 50(5–6), 355–369. <https://doi.org/10.1027/1864-9335/a000395>
- Giacomantonio, M., Jordan, J., Fennis, B. M., & Panno, A. (2014). When the motivational consequences of ego depletion collide: Conservation dominates over reward-seeking. *Journal of Experimental Social Psychology*, 55, 217–220. <https://doi.org/10.1016/j.jesp.2014.07.009>
- Gibson, E. L. (2007). Carbohydrates and mental function: Feeding or impeding the brain? *Nutrition Bulletin*, 32(Suppl. 1), 71–83. <https://doi.org/10.1111/j.1467-3010.2007.00606.x>
- Grahek, I., Musslick, S., & Shenhav, A. (2020). A computational perspective on the roles of affect in cognitive control. *International Journal of Psychophysiology*, 151, 25–34. <https://doi.org/10.1016/j.ijpsycho.2020.02.001>
- Green, L., & Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin*, 130(5), 769–792. <https://doi.org/10.1037/0033-2909.130.5.769>
- Green, L., Myerson, J., Oliveira, L., & Chang, S. E. (2013). Delay discounting of monetary rewards over a wide range of amounts. *Journal of the Experimental Analysis of Behavior*, 100(3), 269–281. <https://doi.org/10.1002/jeab.45>
- Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science*, 7(2), 217–229. <https://doi.org/10.1111/tops.12142>
- Han, S., Lerner, J. S., & Keltner, D. (2007). Feelings and consumer decision making: The appraisal-tendency framework. *Journal of Consumer Psychology*, 17(3), 158–168. [https://doi.org/10.1016/S1057-7408\(07\)70023-2](https://doi.org/10.1016/S1057-7408(07)70023-2)
- Hancock, P. A., & Desmond, P. A. (2001). *Stress, workload, and fatigue*. Lawrence Erlbaum.
- Hartmann, M. N., Hager, O. M., Tobler, P. N., & Kaiser, S. (2013). Parabolic discounting of monetary rewards by physical effort. *Behavioural Processes*, 100, 192–196. <https://doi.org/10.1016/j.beproc.2013.09.014>
- Heckman, J. (1974). Shadow prices, market wages, and labor supply. *Econometrica*, 42(4), 679–694. <https://doi.org/10.2307/1913937>
- Hockey, G. R. J. (2011). A motivational control theory of cognitive fatigue. In P. L. Ackerman (Ed.), *Cognitive fatigue: Multidisciplinary perspectives on current research and future applications* (pp. 167–187). American Psychological Association. <https://doi.org/10.1037/12343-008>
- Hockey, G. R. J., & Earle, F. (2006). Control over the scheduling of simulated office work reduces the impact of workload on mental fatigue and task performance. *Journal of Experimental Psychology: Applied*, 12(1), 50–65. <https://doi.org/10.1037/1076-898X.12.1.50>
- Holtzer, R., Yuan, J., Verghese, J., Mahoney, J. R., Izzetoglu, M., & Wang, C. (2017). Interactions of subjective and objective measures of fatigue defined in the context of brain control of locomotion. *The Journals of Gerontology: Series A*, 72(3), 417–423. <https://doi.org/10.1093/gerona/glw167>
- Hopstaken, J. F., Van Der Linden, D., Bakker, A. B., & Kompier, M. A. (2015). A multifaceted investigation of the link between mental fatigue and task disengagement. *Psychophysiology*, 52(3), 305–315. <https://doi.org/10.1111/psyp.12339>
- Hopstaken, J. F., van der Linden, D., Bakker, A. B., Kompier, M. A., & Leung, Y. K. (2016). Shifts in attention during mental fatigue: Evidence from subjective, behavioral, physiological, and eye-tracking data. *Journal of Experimental Psychology: Human Perception and Performance*, 42(6), 878–889. <https://doi.org/10.1037/xhp0000189>
- Horiuchi, R., Ogasawara, T., & Miki, N. (2018). Fatigue assessment by blink detected with attachable optical sensors of dye-sensitized photovoltaic cells. *Micromachines*, 9(6), Article 310. <https://doi.org/10.3390/mi9060310>
- Hossain, J., Reinish, L., Kayumov, L., Bhuiya, P., & Shapiro, C. (2003). Underlying sleep pathology may cause chronic high fatigue in shift-workers. *Journal of Sleep Research*, 12(3), 223–230. <https://doi.org/10.1046/j.1365-2869.2003.00354.x>
- Jara-Ettinger, J. (2019). Theory of mind as inverse reinforcement learning. *Current Opinion in Behavioral Sciences*, 29, 105–110. <https://doi.org/10.1016/j.cobeha.2019.04.010>
- Jebara, T. (2012). *Machine learning: Discriminative and generative* (Vol. 755). Springer.
- Jha, A. K., Iliff, A. R., Chaoui, A. A., Defossez, S., Bombaugh, M. C., & Miller, Y. R. (2019). *A crisis in health care: A call to action on physician burnout* (Technical Report). Massachusetts Medical Society; Massachusetts Health and Hospital Association; Harvard T.H. Chan School of Public Health; Harvard Global Health Institute.
- Jimura, K., Locke, H. S., & Braver, T. S. (2010). Prefrontal cortex mediation of cognitive enhancement in rewarding motivational contexts. *Proceedings of the National Academy of Sciences*, 107(19), 8871–8876. <https://doi.org/10.1073/pnas.1002007107>
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291. <https://doi.org/10.2307/1914185>
- Karasek, R. A., Jr. (1979). Job demands, job decision latitude, and mental strain: Implications for job redesign. *Administrative Science Quarterly*, 24(2), 285–308. <https://doi.org/10.2307/2392498>
- Kaufman, B. E. (1999). Emotional arousal as a source of bounded rationality. *Journal of Economic Behavior & Organization*, 38(2), 135–144. [https://doi.org/10.1016/S0167-2681\(99\)00002-5](https://doi.org/10.1016/S0167-2681(99)00002-5)
- Keane, M. P., Todd, P. E., & Wolpin, K. I. (2011). The structural estimation of behavioral models: Discrete choice dynamic programming methods and applications. In *Handbook of labor economics* (pp. 331–461). Elsevier. [https://doi.org/10.1016/S0169-7218\(11\)00410-2](https://doi.org/10.1016/S0169-7218(11)00410-2)
- Kelley, N. J., Finley, A. J., & Schmeichel, B. J. (2019). After-effects of self-control: The reward responsivity hypothesis. *Cognitive, Affective, & Behavioral Neuroscience*, 19(3), 600–618. <https://doi.org/10.3758/s13415-019-00694-3>
- Klein-Flügge, M. C., Kennerley, S. W., Saraiva, A. C., Penny, W. D., & Bestmann, S. (2015). Behavioral modeling of human choices reveals dissociable effects of physical effort and temporal delay on reward devaluation. *PLoS Computational Biology*, 11(3), Article e1004116. <https://doi.org/10.1371/journal.pcbi.1004116>
- Knill, D. C., & Pouget, A. (2004). The bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12), 712–719. <https://doi.org/10.1016/j.tins.2004.10.007>
- Kodithuwakku Arachchige, S. N., Burch V, R. F., Chander, H., Turner, A. J., & Knight, A. C. (2021). The use of wearable devices in cognitive fatigue: Current

- trends and future intentions. *Theoretical Issues in Ergonomics Science*. Advance online publication. <https://doi.org/10.1080/1463922X.2021.1965670>
- Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General*, 139(4), 665–682. <https://doi.org/10.1037/a0020198>
- Kuderer, M., Gulati, S., & Burgard, W. (2015). Learning driving styles for autonomous vehicles from demonstration. In *2015 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 2641–2646). Institute of Electrical and Electronics Engineers.
- Kurzban, R. (2010). Does the brain consume additional glucose during self-control tasks? *Evolutionary Psychology: An International Journal of Evolutionary Approaches to Psychology and Behavior*, 8(2), 244–259. <https://doi.org/10.1177/147470491000800208>
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, 112(2), 443–478. <https://doi.org/10.1162/003355397555253>
- Lal, S. K., & Craig, A. (2001). A critical review of the psychophysiology of driver fatigue. *Biological Psychology*, 55(3), 173–194. [https://doi.org/10.1016/S0301-0511\(00\)00085-5](https://doi.org/10.1016/S0301-0511(00)00085-5)
- Laming, D. R. J. (1968). *Information theory of choice-reaction times*. Academic Press.
- Lazarus, R. S., & Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford University Press.
- Le, D., & Provost, E. M. (2013). Emotion recognition from spontaneous speech using hidden markov models with deep belief networks. In *2013 IEEE workshop on automatic speech recognition and understanding* (pp. 216–221). Institute of Electrical and Electronics Engineers.
- Lieder, F., Griffiths, T. L., & Goodman, N. D. (2012). Burn-In, bias, and the rationality of anchoring. In F. Pereira, C. J. C. Burges, L. Bottou & K. Q. Weinberger (Eds.), *NIPS* (pp. 2699–2707). Curran Associates.
- Lieder, F., Shenhav, A., Musslick, S., & Griffiths, T. L. (2018). Rational metareasoning and the plasticity of cognitive control. *PLoS Computational Biology*, 14(4), Article e1006043. <https://doi.org/10.1371/journal.pcbi.1006043>
- Liu, J., Zhang, C., & Zheng, C. (2010). Eeg-based estimation of mental fatigue by using kpca-hmm and complexity parameters. *Biomedical Signal Processing and Control*, 5(2), 124–130. <https://doi.org/10.1016/j.bspc.2010.01.001>
- Loewenstein, G., & Lerner, J. S. (2003). The role of affect in decision making. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 619–642). Oxford University Press.
- Loewenstein, G., & O'Donoghue, T. (2004). *Animal spirits: Affective and deliberative processes in economic behavior*. <https://ssrn.com/abstract=539843>
- Loewenstein, G., Rick, S., & Cohen, J. D. (2008). Neuroeconomics. *Annual Review of Psychology*, 59, 647–672. <https://doi.org/10.1146/annurev.psych.59.103006.093710>
- Luo, H., Lee, P.-A., Clay, I., Jaggi, M., & De Luca, V. (2020). Assessment of fatigue using wearable sensors: A pilot study. *Digital Biomarkers*, 4(1), 59–72. <https://doi.org/10.1159/000512166>
- Luo, Y., & Qi, W. (2019). Pilots' brain cognitive state inference based on remaining life hsmm. In *2019 Chinese Control and Decision Conference (CCDC)* (pp. 5289–5294). Institute of Electrical and Electronics Engineers.
- Magnac, T., & Thesmar, D. (2002). Identifying dynamic discrete decision processes. *Econometrica*, 70(2), 801–816. <https://doi.org/10.1111/1468-0262.00306>
- Maier, S. F., & Seligman, M. E. (1976). Learned helplessness: Theory and evidence. *Journal of Experimental Psychology: General*, 105(1), 3–46. <https://doi.org/10.1037/0096-3445.105.1.3>
- Mairesse, O., & Neu, D. (2016). Tired of blunt tools? sharpening the clinical assessment of fatigue and sleepiness. *Psychiatry Research*, 238, 100–108. <https://doi.org/10.1016/j.psychres.2016.02.005>
- Marx, I., Höpcke, C., Berger, C., Wandschneider, R., & Herpertz, S. C. (2013). The impact of financial reward contingencies on cognitive function profiles in adult adhd. *PLOS ONE*, 8(6), Article e67002. <https://doi.org/10.1371/journal.pone.0067002>
- Massar, S. A., Csathó, Á., & Van der Linden, D. (2018). Quantifying the motivational effects of cognitive fatigue through effort-based decision making. *Frontiers in Psychology*, 9, Article 843. <https://doi.org/10.3389/fpsyg.2018.00843>
- McFadden, D. (1980). Econometric models for probabilistic choice among products. *The Journal of Business*, 53(Suppl. 3), S13–S29. <http://www.jstor.org/stable/2352205>
- Mellers, B. A., Schwartz, A., & Cooke, A. D. (1998). Judgment and decision making. *Annual Review of Psychology*, 49(1), 447–477. <https://doi.org/10.1146/annurev.psych.49.1.447>
- Milyavskaya, M., Inzlicht, M., Johnson, T., & Larson, M. J. (2019). Reward sensitivity following boredom and cognitive effort: A high-powered neurophysiological investigation. *Neuropsychologia*, 123, 159–168. <https://doi.org/10.1016/j.neuropsychologia.2018.03.033>
- Monahan, G. E. (1982). State of the art—A survey of partially observable markov decision processes: Theory, models, and algorithms. *Management Science*, 28(1), 1–16. <https://doi.org/10.1287/mnsc.28.1.1>
- Mor, B., Garhwal, S., & Kumar, A. (2021). A systematic review of hidden markov models and their applications. *Archives of computational methods in engineering*, 28, 1429–1448. <https://doi.org/10.1007/s11831-020-09422-4>
- Morgenstern, O., & Von Neumann, J. (1953). *Theory of games and economic behavior*. Princeton University Press.
- Muraven, M., & Slessareva, E. (2003). Mechanisms of self-control failure: Motivation and limited resources. *Personality and Social Psychology Bulletin*, 29(7), 894–906. <https://doi.org/10.1177/0146167203029007008>
- Musslick, S., Cohen, J. D., & Shenhav, A. (2018). Estimating the costs of cognitive control from task performance: Theoretical validation and potential pitfalls. *CogSci* (pp. 798–803). John Wiley & Sons.
- Musslick, S., Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2015). A computational model of control allocation based on the expected value of control. *The 2nd multidisciplinary conference on reinforcement learning and decision making*.
- Ng, A. Y., & Jordan, M. I. (2002). On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. In T. Dietterich, S. Becker & Z. Ghahramani (Eds.), *Advances in neural information processing systems* (pp. 841–848). MIT Press.
- Ng, A. Y., & Russell, S. J. (2000). Algorithms for inverse reinforcement learning. In P. Langley (Ed.), *Proceedings of the seventeenth international conference on machine learning* (pp. 663–670). Morgan Kaufmann Publishers.
- Norets, A., & Tang, X. (2014). Semiparametric inference in dynamic binary choice models. *Review of Economic Studies*, 81(3), 1229–1262. <https://doi.org/10.1093/restud/rdt050>
- Obermaier, B., Guger, C., Neuper, C., & Pfurtscheller, G. (2001). Hidden markov models for online classification of single trial eeg data. *Pattern Recognition Letters*, 22(12), 1299–1309. [https://doi.org/10.1016/S0167-8655\(01\)00075-7](https://doi.org/10.1016/S0167-8655(01)00075-7)
- Otto, A. R., Skatova, A., Madlon-Kay, S., & Daw, N. D. (2015). Cognitive control predicts use of model-based reinforcement learning. *Journal of Cognitive Neuroscience*, 27(2), 319–333. https://doi.org/10.1162/jocn_a_00709
- Padmala, S., & Pessoa, L. (2010). Interactions between cognition and motivation during response inhibition. *Neuropsychologia*, 48(2), 558–565. <https://doi.org/10.1016/j.neuropsychologia.2009.10.017>
- Padoa-Schioppa, C. (2011). Neurobiology of economic choice: A good-based model. *Annual Review of Neuroscience*, 34, 333–359. <https://doi.org/10.1146/annurev-neuro-061010-113648>
- Pattyn, N., Neyt, X., Henderickx, D., & Soetens, E. (2008). Psychophysiological investigation of vigilance decrement: Boredom or cognitive fatigue? *Physiology & Behavior*, 93(1–2), 369–378. <https://doi.org/10.1016/j.physbeh.2007.09.016>
- Pattyn, N., Van Cutsem, J., Dessy, E., & Mairesse, O. (2018). Bridging exercise science, cognitive psychology, and medical practice: Is “cognitive fatigue” a remake of “the emperor’s new clothes”? *Frontiers in Psychology*, 9, Article 1246. <https://doi.org/10.3389/fpsyg.2018.01246>
- Pavlidis, I., Dcosta, M., Taamneh, S., Manser, M., Ferris, T., Wunderlich, R., Akleman, E., & Tsiamyrtzis, P. (2016). Dissecting driver behaviors under cognitive, emotional, sensorimotor, and mixed stressors. *Scientific Reports*, 6(1), 1–12. <https://doi.org/10.1038/srep25651>

- Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, 9(2), 148–158. <https://doi.org/10.1038/nrn2317>
- Pine, A., Seymour, B., Roiser, J. P., Bossaerts, P., Friston, K. J., Curran, H. V., & Dolan, R. J. (2009). Encoding of marginal utility across time in the human brain. *Journal of Neuroscience*, 29(30), 9575–9581. <https://doi.org/10.1523/JNEUROSCI.1126-09.2009>
- Power, S. D., Falk, T. H., & Chau, T. (2010). Classification of prefrontal activity due to mental arithmetic and music imagery using hidden markov models and frequency domain near-infrared spectroscopy. *Journal of Neural Engineering*, 7(2), Article 026002. <https://doi.org/10.1088/1741-2560/7/2/026002>
- Prévost, C., Pessiglione, M., Météreau, E., Cléry-Melin, M.-L., & Dreher, J.-C. (2010). Separate valuation subsystems for delay and effort decision costs. *Journal of Neuroscience*, 30(42), 14080–14090. <https://doi.org/10.1523/JNEUROSCI.2752-10.2010>
- Puranam, P., Stieglitz, N., Osman, M., & Pillutla, M. M. (2015). Modelling bounded rationality in organizations: Progress and prospects. *Academy of Management Annals*, 9(1), 337–392. <https://doi.org/10.5465/19416520.2015.1024498>
- Puterman, M. L. (2014). *Markov decision processes: Discrete stochastic dynamic programming*. Wiley.
- Raichle, M. E., & Mintun, M. A. (2006). Brain work and brain imaging. *Annual Review of Neuroscience*, 29, 449–476. <https://doi.org/10.1146/annurev.neuro.29.051605.112819>
- Razmpa, E., Niat, K. S., & Saedi, B. (2011). Urban bus drivers' sleep problems and crash accidents. *Indian Journal of Otolaryngology and Head & Neck Surgery*, 63(3), 269–273. <https://doi.org/10.1007/s12070-011-0235-5>
- Reinhardt, L., & Wahba, M. A. (1975). Expectancy theory as a predictor of work motivation, effort expenditure, and job performance. *Academy of Management Journal*, 18(3), 520–537. <https://doi.org/10.2307/255682>
- Ross, S. M. (2014). *Introduction to stochastic dynamic programming*. Academic Press.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, 80(1), 1–28. <https://doi.org/10.1037/h0092976>
- Rust, J. (1987). Optimal replacement of gmc bus engines: An empirical model of Harold zurcher. *Econometrica*, 55(5), 999–1033. <http://www.jstor.org/stable/1911259>
- Rust, J. (1994). Structural estimation of markov decision processes. *Handbook of econometrics* (Vol. 4, pp. 3081–3143). Elsevier. [https://doi.org/10.1016/S1573-4412\(05\)80020-0](https://doi.org/10.1016/S1573-4412(05)80020-0)
- Samuelson, P. A. (1938). A note on the pure theory of consumer's behaviour. *Economica*, 5(17), 61–71. <https://doi.org/10.2307/2548836>
- Sandry, J., Genova, H. M., Dobryakova, E., DeLuca, J., & Wylie, G. (2014). Subjective cognitive fatigue in multiple sclerosis depends on task length. *Frontiers in Neurology*, 5, Article 214. <https://doi.org/10.3389/fneur.2014.00214>
- Schmidt, L., Lebreton, M., Cléry-Melin, M.-L., Daunizeau, J., & Pessiglione, M. (2012). Neural mechanisms underlying motivation of mental versus physical effort. *PLoS Biology*, 10(2), Article e1001266. <https://doi.org/10.1371/journal.pbio.1001266>
- Schmitt, F., Bieg, H.-J., Herman, M., & Rothkopf, C. A. (2017). I see what you see: Inferring sensor and policy models of human real-world motor behavior. In S. Singh & S. Markovitch (Chairs, *Proceedings of the thirty-first AAAI conference on artificial intelligence* (pp. 3797–3803). AAAI Press.
- Seymour, B., O'Doherty, J. P., Dayan, P., Koltzenburg, M., Jones, A. K., Dolan, R. J., Friston, K. J., & Frackowiak, R. S. (2004). Temporal difference models describe higher-order learning in humans. *Nature*, 429(6992), 664–667. <https://doi.org/10.1038/nature02581>
- Shen, J., Barbera, J., & Shapiro, C. M. (2006). Distinguishing sleepiness and fatigue: Focus on definition and measurement. *Sleep Medicine Reviews*, 10(1), 63–76. <https://doi.org/10.1016/j.smrv.2005.05.004>
- Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*, 79(2), 217–240. <https://doi.org/10.1016/j.neuron.2013.07.007>
- Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*, 40, 99–124. <https://doi.org/10.1146/annurev-neuro-072116-031526>
- Smallwood, R. D., & Sondik, E. J. (1973). The optimal control of partially observable markov processes over a finite horizon. *Operations Research*, 21(5), 1071–1088. <https://doi.org/10.1287/opre.21.5.1071>
- Sohn, J.-H., Kim, H.-E., Sohn, S., Seok, J.-W., Choi, D., & Watanuki, S. (2015). Effect of emotional arousal on inter-temporal decision-making: An fmri study. *Journal of Physiological Anthropology*, 34(1), 1–8. <https://doi.org/10.1186/s40101-015-0047-5>
- Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economics Letters*, 8(3), 201–207. [https://doi.org/10.1016/0165-1765\(81\)90067-7](https://doi.org/10.1016/0165-1765(81)90067-7)
- Thormdike, E. (1900). Mental fatigue. I. *Psychological Review*, 7(6), 547–579. <https://doi.org/10.1037/h0069511>
- Todd, P. E., & Wolpin, K. I. (2006). Assessing the impact of a school subsidy program in mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility. *American Economic Review*, 96(5), 1384–1417. <https://doi.org/10.1257/aer.96.5.1384>
- Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76(1), 31–48. <https://doi.org/10.1037/h0026750>
- Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, 114(1), 104–132. <https://doi.org/10.1037/0033-295X.114.1.104>
- van Steenbergen, H., Booij, L., Band, G. P., Hommel, B., & van der Does, A. W. (2012). Affective regulation of cognitive-control adjustments in remitted depressive patients after acute tryptophan depletion. *Cognitive, Affective, & Behavioral Neuroscience*, 12(2), 280–286. <https://doi.org/10.3758/s13415-011-0078-2>
- Völker, I., Kirchner, C., & Bock, O. L. (2016). On the relationship between subjective and objective measures of fatigue. *Ergonomics*, 59(9), 1259–1263. <https://doi.org/10.1080/00140139.2015.1110622>
- Vul, E., Goodman, N., Griffiths, T. L., & Tenenbaum, J. B. (2014). One and done? Optimal decisions from very few samples. *Cognitive Science*, 38(4), 599–637. <https://doi.org/10.1111/cogs.12101>
- Wabba, M. A., & House, R. J. (1974). Expectancy theory in work and motivation: Some logical and methodological issues. *Human Relations*, 27(2), 121–147. <https://doi.org/10.1177/001872677402700202>
- Wall, T. D., Jackson, P. R., Mullarkey, S., & Parker, S. K. (1996). The demands—control model of job strain: A more specific test. *Journal of Occupational and Organizational Psychology*, 69(2), 153–166. <https://doi.org/10.1111/j.2044-8325.1996.tb00607.x>
- Wang, C., Alvarez, S., Ruiz, C., & Moonis, M. (2013). Computational modeling of sleep stage dynamics using weibull semi-markov chains. In D. Stacey, J. Solé-Casals, A. Fre, & H. Gamboa (Eds.), *HEALTHINF 2013—Proceedings of the international conference on health informatics* (pp. 122–130). SciTePress.
- Westbrook, A., & Braver, T. S. (2015). Cognitive effort: A neuroeconomic approach. *Cognitive, Affective, & Behavioral Neuroscience*, 15(2), 395–415. <https://doi.org/10.3758/s13415-015-0334-y>
- Westbrook, A., Kester, D., & Braver, T. S. (2013). What is the subjective cost of cognitive effort? load, trait, and aging effects revealed by economic preference. *PLOS ONE*, 8(7), Article e68210. <https://doi.org/10.1371/journal.pone.0068210>
- Wylie, G., Genova, H., DeLuca, J., & Dobryakova, E. (2017). The relationship between outcome prediction and cognitive fatigue: A convergence of paradigms. *Cognitive, Affective, & Behavioral Neuroscience*, 17(4), 838–849. <https://doi.org/10.3758/s13415-017-0515-y>
- Wylie, G. R., Yao, B., Sandry, J., & DeLuca, J. (2020). Using signal detection theory to better understand cognitive fatigue. *Frontiers in Psychology*, 11, Article 579188. <https://doi.org/10.3389/fpsyg.2020.579188>
- Yu, S.-Z. (2010). Hidden semi-markov models. *Artificial Intelligence*, 174(2), 215–243. <https://doi.org/10.1016/j.artint.2009.11.011>
- Zhang, C., Zheng, C., & Yu, X. (2008). Evaluation of mental fatigue based on multiphysiological parameters and kernel learning algorithms. *Chinese Science Bulletin*, 53(12), 1835–1847. <https://doi.org/10.1007/s11434-008-0245-1>

Appendix A

Detailed Belief Updating Process

Here we provide more details of the Bayesian belief update. Let \mathbb{X} be the (finite dimensional) probability simplex in $\mathbb{R}^{|\mathbb{S} \times \mathbb{D}|}$ and $x_t \triangleq \{P(s_t, \tau_t | \zeta_t), \forall (s_t, \tau_t) \in \mathbb{S} \times \mathbb{D}\} \in \mathbb{X}$ be the Bayesian belief, which is the posterior distribution over possible fatigue states and its duration times (s_t, τ_t) given history ζ_t . Then, $\forall t, x_t \in \mathbb{X}$, the size of belief x_t remains constant and does not grow to infinity. To see this, for a given belief x_t at time t —which is a matrix of size $S \times D$ (S rows and D columns)—the updated (posterior) belief x_{t+1} at time $t+1$ also has the same size as x_t , where the belief update is given by:

$$\begin{aligned} x_{t+1}(s, \tau) &= P(s_{t+1} = s, \tau_{t+1} = \tau | z_{t+1}, a_t, \zeta_t) \\ &= \frac{\sum_{s_t, \tau_t} P(s_{t+1} = s, \tau_{t+1} = \tau, z_{t+1} | s_t, \tau_t, a_t) x_t(s_t, \tau_t)}{\sigma(x_t, a_t, z_{t+1})} \\ &= \frac{P(s_{t+1} = s, \tau_{t+1} = \tau, z_{t+1} | x_t, a_t)}{\sigma(x_t, a_t, z_{t+1})}, \end{aligned} \quad (\text{A1})$$

where the denominator is $\sigma(x_t, a_t, z_{t+1}) \triangleq \sum_{s_{t+1}, \tau_{t+1}} \sum_{s_t, \tau_t} P(s_{t+1}, \tau_{t+1}, z_{t+1} | s_t, \tau_t, a_t) x_t(s_t, \tau_t) = P(z_{t+1} | x_t, a_t)$ is called the reduced-form observation probability (Chang et al., 2020). We denote this updating procedure by $x_{t+1} = \lambda(x_t, a_t, z_{t+1})$. The

numerator of Equation A1 can be further factorized by considering $\tau_t = 0$ and $\tau_t > 0$:

$$\begin{aligned} &P(s_{t+1} = s, \tau_{t+1} = \tau, z_{t+1} | x_t, a_t) \\ &= \underbrace{\sum_{s_t \neq s} P(s_{t+1} = s, \tau_{t+1} = \tau, z_{t+1} | s_t, \tau_t, a_t) x(s_t, \tau_t)}_{\tau_t = 0} \\ &+ \underbrace{P(s_{t+1} = s, \tau_{t+1} = \tau, z_{t+1} | s_t = s, \tau_t, a_t) x(s_t = s, \tau_t)}_{\tau_t = \tau + 1}. \end{aligned}$$

The updated belief relies on the current state s_t and time left τ_t in s_t . If $\tau_t = 0$, then the current fatigue state $s_t = s$ transitions to a different state $s_{t+1} = s' (s' \neq s)$, generating a new sojourn time τ_{t+1} and an observation z_{t+1} according to the environment state dynamics. If $\tau_t > 0$, then the current fatigue state persists ($s_{t+1} = s_t$). To find the probability associated $\tau_{t+1} = \tau$, we use the probability associated with $\tau_t = \tau + 1$ because $\tau_{t+1} = \tau_t - 1$ when $\tau_t > 0$. The σ function in the Equation A1 serves as a regulator for the sum of all the elements of x_{t+1} equals 1.

Appendix B

Robustness Analysis Experiment of the Number of Fatigue States

We designed a numerical experiment and generated data assuming three true fatigue states: State 1 (*not fatigued* state), State 2 (*medium fatigue* state), and State 3 (*fatigued* state). Assuming the number of fatigue states is not known in advance, we started estimation from two fatigue states, then three fatigue states, and so on.

Specifically, we began with the “ground truth” model to generate synthetic data and then used this data set to recover the model parameters wherein the number of fatigue states was an unknown parameter.

1. The state transition probabilities are shown in Table B1. Note that we still assume that easier tasks lead to faster recovery as in section Simulated N -back Tasks with Changing Fatigue. When choosing and performing $a = 1$ (1-back) task, the fatigue state will improve. When choosing and performing $a = 2, 3$ (2, 3-back) task, fatigue state will deteriorate, and harder tasks lead to a higher probability of deterioration. For example, in Table B1(a), when the current state is either 1 or 2, and the choice is to select 1 (1-back, easiest), the state transition can only lead to a better state (not fatigued). And when the current state is 3, choosing and performing the 1-back task leads to less fatigued state (state 1, 2). Similar rules apply in Table B1(b–c).
2. The observation probabilities $P(z|s)$ are shown in Table B2.

3. Sojourn time distribution parameters are given by Table B3, wherein the sojourn time distribution takes the form $P(\tau_{t+1} | s_{t+1}, a_t, \tau_t = 0)$. For illustration purpose, we assumed that sojourn times follow Binomial distributions with two parameters: a probability parameter and the total number of trials, the latter of which is fixed to 4 (according to experiment design, not to be estimated). Table B3 summarizes the parameter values used to define

Table B1

State Transition Probabilities for Different Choices

State	$s' = 1$	$s' = 2$	$s' = 3$
a. State transition probabilities when the choice is 1-back task, $P(s' s, a = 1, \tau = 0)$.			
$s = 1$	1	0	0
$s = 2$	1	0	0
$s = 3$	0.4 (0.40)	0.6 (0.60)	0
b. State transition probabilities when the choice is 2-back task, $P(s' s, a = 2, \tau = 0)$.			
$s = 1$	0	0.9 (0.86)	0.1 (0.14)
$s = 2$	0	0	1
$s = 3$	0	0	1
c. State transition probabilities when the choice is 3-back, $P(s' s, a = 3, \tau = 0)$.			
$s = 1$	0	0.7 (0.70)	0.3 (0.30)
$s = 2$	0	0	1
$s = 3$	0	0	1

Note. Values in the parentheses are the estimates from the 3-state model.

(Appendices continue)

Table B2*Observation Probabilities $P(z|s)$*

State	$z = 1$	$z = 2$	$z = 3$	$z = 4$
$s = 1$	0.1 (0.10)	0.2 (0.20)	0.3 (0.30)	0.4 (0.40)
$s = 2$	0.2 (0.19)	0.3 (0.29)	0.4 (0.41)	0.1 (0.11)
$s = 3$	0.4 (0.41)	0.3 (0.30)	0.2 (0.20)	0.1 (0.09)

Note. Values in the parentheses are the estimates from the 3-state model.

those Binomial distributions for each choice a_t and each future state s_{t+1} .

- External (performance dependent) monetary rewards are the same as used in the article.

$$\text{incentive}(a; c_p) = \begin{cases} 1 & a = 1, \\ 3 & a = 2, \\ 6 & a = 3, \end{cases}$$

where c_p denotes the (predefined) performance criteria required to receive incentives.

Appendix C

Proofs

Proof of Theorem 1

The proof is by induction as in Chang et al. (2020) and Smallwood and Sondik (1973). Let $\zeta_{t+1} = \zeta'$, $\varepsilon_{t+1} = \varepsilon'$. Assume $g_{t+1}(\zeta', \varepsilon') = V_{t+1}(x', \varepsilon')$, then at time t , we have:

$$\begin{aligned} g_t(\zeta, \varepsilon) &= \max_a \left\{ \sum_{s, \tau} P(s, \tau | \zeta) r(s, a) + \varepsilon(a) \right. \\ &\quad \left. + \beta \sum_{z'} P(z' | \zeta, a) \int V_{t+1}(x', \varepsilon') d\mu(\varepsilon') \right\} \\ &= \max_a \left\{ r(x, a) + \varepsilon(a) + \beta \sum_{z'} \sigma(x, a, z') \int V_{t+1}(\lambda(x, a, z'), \varepsilon') d\mu(\varepsilon') \right\}, \\ &= V_t(x, \varepsilon). \end{aligned}$$

where the second equation is by $P(z' | \zeta, a) = \sigma(x, a, z')$ and $x' = \lambda(x, a, z')$. \square

Identification of Fatigue Dynamics

We explain the identification of the dynamics of fatigue states and their duration times as follows. Such dynamics are defined by $\{P_{\theta_D}(s', \tau', z | s, \tau, a) : s, s' \in \mathbb{S}, \tau, \tau' \in \mathbb{D}, a \in A\}$ (parameterized by θ_D). Both reduced-form observation probability and updated Bayesian belief depend on P_{θ_D} and are denoted as $\sigma_{\theta_D}, x_{t, \theta_D}$, respectively.

Given that:

$$\underbrace{\sigma_{\theta_D}(z_{t+1}, x_{t, \theta_D}, a_t)}_{\text{a function of } \theta_D} = \underbrace{\hat{P}(z_{t+1} | \zeta_t, a_t)}_{\text{from data}},$$

where $\hat{P}(z_{t+1} | \zeta_t, a_t)$ is a function of data, one would expect that θ_D could be identified by matching σ_{θ_D} with empirical probabilities

5. One-step subjective value parameters

$$\theta_3 = \{\theta_{3,1}, \theta_{3,2}\} = \{2.5, 0.5\}.$$

We used the aforementioned model parameters to generate a sample data set containing 2,000 sample paths with 80 decision epochs for each sample path. Then we fit the models assuming 2-, 3-, and 4-state settings, respectively.

Table B3*Sojourn Time Parameters*

Choice/Action	$s_{t+1} = 1$	$s_{t+1} = 2$	$s_{t+1} = 3$
$a_t = 1$	0.3 (0.29)	0.8 (0.83)	—
$a_t = 2$	—	0.7 (0.69)	0.2 (0.20)
$a_t = 3$	—	0.4 (0.38)	0.4 (0.41)

Note. Values in the parentheses are the estimates of the 3-state model. Table cells with no value indicates those scenarios will not occur.

$\hat{P}(z_{t+1} | \zeta_t, a_t)$. To test this idea, consider the first-period data $\{x_0, a_0, z_1\}$:

$$\begin{aligned} \sigma_{\theta_D}(z_1, x_0, a_0) &= \sum_{s_1, \tau_1} \sum_{s_0, \tau_0} x_0(s_0, \tau_0) P_{\theta_D}(s_1, \tau_1, z_1 | s_0, \tau_0, a_0) \\ &= \sum_{s_0, \tau_0} x_0(s_0, \tau_0) P_{\theta_D}(z_1 | s_0, \tau_0, a_0) \\ &= \hat{P}(z_1 | x_0, a_0) \end{aligned} \quad (C1)$$

The issue is that Equation C1 is not enough to uniquely determine $P(s_{t+1}, \tau_{t+1}, z_{t+1} | s_t, \tau_t, a_t)$: It is possible that there exist $\theta_D \neq \theta'_D$ and $(s_t + 1, \tau_t + 1)$ such that:

$$P_{\theta_D}(s_{t+1}, \tau_{t+1}, z_{t+1} | s_t, \tau_t, a_t) \neq P_{\theta'_D}(s_{t+1}, \tau_{t+1}, z_{t+1} | s_t, \tau_t, a_t)$$

but

$$P_{\theta_D}(z_{t+1} | s_t, \tau_t, a_t) = P_{\theta'_D}(z_{t+1} | s_t, \tau_t, a_t).$$

In this case, consider two periods of data $\{x_0, a_0, z_1, a_1, z_2\}$. We show that for two dynamics of fatigue θ_D, θ'_D :

$$\theta'_D = \theta_D \Leftrightarrow x_{1, \theta_D} = x_{1, \theta'_D}, \quad (C2)$$

where “ \Leftrightarrow ” stands for “if and only if.”

“ \Rightarrow ”: is trivial to see.

“ \Leftarrow ”: Since

$$x_{1, \theta_D}(s, \tau) = \frac{\sum_{s_0, \tau_0} P_{\theta_D}(s_1 = s, \tau_1 = \tau, z_1 | s_0, \tau_0, a_0) x_0(s_0, \tau_0)}{\sigma_{\theta_D}(x_0, a_0, z_1)},$$

where x_0 is given and $\sigma_{\theta_D}(x_0, a_0, z_1) = \sigma_{\theta'_D}(x_0, a_0, z_1) = \hat{P}(z_1 | \zeta_0, a_0), \forall x_0, a_0, z_1$. Thus, if $x_{1, \theta_D} = x_{1, \theta'_D}$, then $\theta_D = \theta'_D$.

(Appendices continue)

Furthermore, we claim that:

$$x_{1,\theta_D} = x_{1,\theta'_D} \Leftrightarrow \sigma_{\theta_D}(x_{1,\theta_D}, a_1, z_2) = \sigma_{\theta'_D}(x_{1,\theta'_D}, a_1, z_2). \quad (C3)$$

“ \Rightarrow ”: holds because $x_{1,\theta_D} = x_{1,\theta'_D} \Rightarrow \theta_D = \theta'_D$ (Equation C2).

“ \Leftarrow ”: holds because

$$\begin{aligned} \sigma(x_1, a_1, z_2) &= \sum_{s_2, \tau_2, s_1, \tau_1} P(s_2, \tau_2, z_2 | s_1, \tau_1, a_1) x_1(s_1, \tau_1) \\ &= \sum_{s_1, \tau_1} P(z_2 | s_1, \tau_1, a_1) x_1(s_1, \tau_1). \end{aligned}$$

Note that $P_{\theta_D}(z_2 | s_1, \tau_1, a_1) = P_{\theta'_D}(z_2 | s_1, \tau_1, a_1)$ is guaranteed by first period of data. If the second period of data is also matched, that is,

$$\sigma_{\theta_D}(z_2, x_{1,\theta_D}, a_1) = \sigma_{\theta'_D}(z_2, x_{1,\theta'_D}, a_1) = \hat{P}(z_2 | \zeta_1, a_1),$$

then $x_{1,\theta_D} = x_{1,\theta'_D}$.

Hence, by matching two periods of data, Equations C2 and C3 hold and $P_{\theta_D}(s', \tau' | s, \tau, a)$ can be uniquely determined, or identified.

Identification of One-Step Subjective Values

We illustrate the identification of one-step subjective values in the case of ε being multivariate *i.i.d* Gumbel. Hence the CCP has the closed form:

$$\pi(a|x) = \frac{\exp Q(x, a)}{\sum_{a'} \exp Q(x, a')}.$$

Consequently, we have the mapping from the difference in $Q(x, a)$ (compared to a reference action a^0) to the difference in CCP (compared to a reference action a^0):

$$\frac{\pi(a|x)}{\pi(a^0|x)} = \exp(Q(x, a) - Q(x, a^0)). \quad (C4)$$

Note that:

$$\begin{aligned} Q(x, a^0) &= r(x, a^0) + \beta \sum_{z'} \sigma(x, a^0, z') \log \sum_{a'} \exp Q(\lambda(x, a, z'), a') \\ &= r(x, a^0) + \beta \sum_{z'} \sigma(x, a^0, z') Q(\lambda(x, a, z'), a^0) \\ &\quad + \beta \sum_{z'} \sigma(x, a^0, z') \log \sum_{a'} \exp(Q(\lambda(x, a, z'), a') - Q(\lambda(x, a, z'), a^0)), \end{aligned}$$

by the fact that $\log \sum_{i=1}^N \exp(x_i) = c + \log \sum_{i=1}^N \exp(x_i - c)$.

Since the dynamics of fatigue state and its duration are identifiable, $\{x_t\}_{t > 0}$ can be calculated for each sample path and $\pi(a|x)$ is obtainable from the data. Consequently,

$$\begin{aligned} &\sum_{z'} \sigma(x, a^0, z') \log \sum_{a'} \exp(Q(\lambda(x, a, z'), a') - Q(\lambda(x, a, z'), a^0)) \\ &= \sum_{z'} \sigma(x, a^0, z') \sum_{a'} \frac{\pi(a' | \lambda(x, a, z'))}{\pi(a^0 | \lambda(x, a, z'))} \end{aligned}$$

can be estimated from data (note $x' = \lambda(x, a, z')$). Since the subjective values for a reference choice $\{r(s, a^0), s \in \mathbb{S}\}$ are known, the expected cumulative subjective value function $Q(x, a^0)$, $x \in \mathbb{X}$ can be uniquely determined by contraction mapping theorem [and so do $Q(x, a)$, $a \neq a^0$, by using Equation C4]. To find subsection values r , we simply use:

$$r(x, a) = Q(x, a) - \beta \sum_{z'} \sigma(x, a, z') \log \sum_{a'} \exp Q(\lambda(x, a, z'), a'),$$

where $r(x, a) = \sum_s x(s) r(s, a)$. Hence, $r(s, a)$ can be identified by solving a set of equations containing different beliefs. \square

Received June 22, 2021

Revision received December 3, 2021

Accepted February 15, 2022 ■