Time allocation and meta-cognition:

A computational approach towards the organization of motivation

A Dissertation

SUBMITTED TO THE FACULTY OF THE

UNIVERSITY OF MINNESOTA

BY

Dominic Mussack

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Dr. Paul Schrater, Advisor

Dedicated to Ashley

Acknowledgements

It is hard to consider how much I owe to the people in my life.

First, I would like to thank my advisor, Paul Schrater, for seriously considering me as a scholar, driving me to never stop learning, and for genuinely caring about my growth as a researcher and person.

I also would like to thank my committee members: Dan Kersten, for both broad and deep talks; Wilma Koutstaal, for helping me start research, and Dave Stephens, for making hard things clear. I owe you all a debt for your patience and intellect.

For those who helped with data collection over these years, in particular Sewon Oh, Lilly Polivoda, Hannah Schewe, Aaron Sattler, Windy Therior, and Ganesh Rakate. Special thanks also for Matthias Böhmer, for generously providing his data for analysis. I want to thank the Multi-Sensory Perception Lab within the Center for Applied and Translational Sensory Science (CATSS), in particular Andrew Byrne, for their shared research, equipment, and help with data collection. I also want to acknowledge the Minnesota Supercomputing Institute (MSI) at the University of Minnesota for providing computational and data storage resources.

To the members of the Schrater Lab (CoMoCo), for helping me bridge multiple fields at once. Also I want to thank the members of the xCIT lab for helping me bridge countries.

To my fellow graduate students and friends, in particular Windy Therior, Thomas Christie, Brent Carpenter, and Robert Edge. You have been invaluable collaborators and friends, and I can't thank you enough for all we've done together. You've made graduate school bearable.

To my family, for being home in so many ways. I love you all.

Finally, to Ashley Clayson. There are no words good enough.

Abstract

How do people allocate time and effort across tasks? This dissertation takes a computational psychology perspective, and puts forward the theory that motivation solves the meta-cognitive problem of allocating resources to different tasks by computing task priority. Motivation research has previously distinguished between two dissociable components of motivation: directing and energizing. These two components refer to different resources that must be allocated: time and effort. We explore the way humans allocate effort by taking advantage of simple decision making tasks and manipulating either task or background information. We develop a novel method that allows researchers to integrate an array of biometrics that capture how decision processes are modulated. We then extend work from optimal foraging theory to account for human tasks in order to analyze how humans allocate time. We derive various results that match time use behavior across domains. Finally, we apply the structural implications of this theory to make predictions in large scale time use data sets. Humans must often schedule mutually exclusive goals to fulfill mutually exclusive needs, which requires us to allocate time across tasks via a priority computation. Re-framing motivation from a resource allocation perspective, and highlighting the unique problem of time allocation, has implications across human decision making behavior, and we demonstrate its relevance in multiple domains.

Contents

List of Figures

Chapter 1

Introduction

> "Most people can motivate themselves to do things simply by knowing that those things need to be done. But not me. For me, motivation is this horrible, scary game where I try to make myself do something while I actively avoid doing it."
>
> – Allie Brosh, *Hyperbole and a Half*

In the webcomic Hyperbole and a Half, Allie Brosh describes some common experiences for those with depression (Brosh, 2013). In playing "The Game of Motivation," she experiences the difficulty of performing otherwise simple tasks, such as returning a videotape to a rental store. This difficulty has nothing to do with being unable to perform the actions involved in the task. It is neither a physically nor cognitively demanding task, and she eventually performs it quite easily. This difficulty she experiences is motivational; the difficulty to *engage in the task itself*.

This dissertation frames the concept of *motivation* from the meta-cognitive perspective of allocating resources across tasks. Motivation more generally refers to the underlying cause or reason for overt behavior (Niv, Joel, & Dayan, 2006; Bolles, 1975; Berridge, 2004; Carver & Scheier, 1998). In other words, motivation determines what we do and when (Simon, 1967; Collier & Rovee-Collier, 1983). Allocating time well has been essential for survival and success across species, and this suggests there must be a core set of principles that organize this behavior. *Task engagement* is a manifestation of these motivational processes, and it is one of the core phenomena that motivation seeks to explain. This phenomenon represents selecting one task over another to engage in, which makes it a kind of decision problem. However, unlike simple decisions, task engagement involves solving a set of problems which are *meta-cognitive* (Cox, Oates, & Perlis, 2011), in that it determines which task to allocate resources to and when, rather than how to solve a given task.

This dissertation examines human engagement and time allocation through analysis of phenomenon, mathematical theory, and both experimental and large-scale data sets. It is meant to be a series of self-contained chapters, with short connections to bridge the overall work. Chapter 2 provides a theoretical framework for how motivation impacts task engagement, drawing on optimal scheduling theory to explain how people schedule goals to satisfy needs. Scheduling theory suggests that humans solve this problem through a priority queue that allocates time and resources to tasks based on their priority. This priority can be thought of as the motivational impetus to engage in a task, and various motivational effects can be re-conceptualized by understanding how informative cues impact priority computations.

Within motivation research, a common distinction is made between motivation's *directing* effects, which task people select and when, and *energizing* effects, the effort or vigor applied towards these tasks (Niv, Joel, & Dayan, 2006; Dickinson & Balleine, 2002). This distinction can also be viewed from how a task benefits from different resources of time or effort (see Figure 1.1). When tasks are mutually exclusive, they require time allocated to them but might not benefit from more effort than some minimum. By contrast, some tasks might significantly change depending on how effort is allocated. This dissertation explores both components of motivation, and the different implications of how effort (in Chapters 3 and 4) or time (in Chapters 5 and 6) are allocated.

The ability to impact resource allocation in decision making requires an interface be-

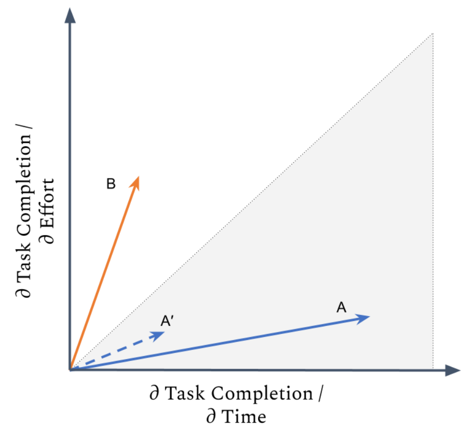

Figure 1.1: Time allocation versus effort allocation in tasks, including the two common components of directing and energizing aspects of motivation. The horizontal axis is the change in task completion given change in time allocation, while the vertical axis is change in task completion given change in effort allocation to the task. Each vector represents how an individual task's probability of completion is impacted by change in effort and time, or how much the task is benefited by adding more effort or time to it.

tween decision circuits and a modulatory process that can set the resources allocated to the circuit (and therefore task). Chapter 3 provides a methodology to investigate effort allocation in a decision making task by providing a measure of modulatory processes. We look for evidence that we can modulate the amount of effort in a simple task by adding background distractors and measuring a large set of redundant biophysical measures (including EEG, heart rate, pupilometry, and galvanic skin response). We use directed dimensionality reduction combined with a simple model of decision processes (the drift diffusion model) to extract the decision-relevant components of these biometrics, allowing us to measure the interface between decision processes and higher level modulatory processes.

We further explore how these decision processes are controlled by modulation in Chapter 4, which investigates how confidence reflects task-relevant information reliability, and allows adaptive changes in the decision process. We investigate information integration by using a system identification approach in a simple decision task, modifying the amount of information available within a trial. We use this method to infer the relative impact of the timing of information as it relates to subjective confidence and performance. We find that integration time is adaptively set in response to forecasted information reliability, and this reliability is reflected in confidence.

We then investigate how time allocation, as an optimal stopping problem, structurally shapes human task switching behavior in Chapter 5. We do this by extending optimal foraging theory to more generic human tasks, specifically those tasks with satisfaction constraints and those motivated by intrinsic factors. The optimal solution allows us to show how background and foreground factors impact time allocation, explaining a large range of self-interruption and spontaneous task switching phenomena.

Chapter 6 takes the formal results from Chapter 5 and applies them to two different time usage data sets, a mobile phone application switching data set and the American Time Use Survey. The theory makes structural predictions on the relationships between background and foreground time use. We demonstrate how a simple measure of background context, that is, nearby activities, can predict foreground activity's time.

Finally, we conclude with a discussion of the future impact of this work, including extensions and implications. By viewing task engagement as the result of an optimal resource allocation problem, we can frame otherwise disparate motivational phenomenon under a single framework.

Much of this dissertation was the result of collaborations between myself and others.

Chapter 3 represents work between Windy Therior and myself. We both designed the study and experiment, collected data, performed most preprocessing, and wrote the content. Windy created most of the explanatory figures (which we designed together), and she was responsible for gathering the hardware and software necessary to conduct biometric data collection (which was no small task). She also preformed the model fits for the drift diffusion model. I coded up the experiment, performed the partial least squares analysis, the heart rate and EEG preprocessing (specifically the rate filtering and temporal basis functions), and the cross validation. Chapter 4 also represents work between Windy Therior, myself, and Hannah Schewe. We all designed the study and experiment and collected data together. Hannah initially designed and ran the logistic regression analysis, which I made slight modifications to. Windy performed all basic visualizations, and we both wrote the content. I designed and ran the hazard analysis, coded the experiment, and did the background review.

Chapter 5 and Chapter 6 represent work between myself and Robert Edge. We both contributed to theory and wrote the initial content and designed the analysis for Chapter 6. Robert ran the initial analysis on the mobile phone data set for Chapter 6, while I replicated those results in the time use survey. Robert was also responsible for the empowerment calculations, while I was responsible for background review, and computations on deadlines and uncertainty.

Chapter 2

Motivation, Engagement, and Time Allocation

2.1 Introduction

Consider the course of engagement involved in writing (see Figure 2.1). A writer might plan to write at 9:00 am that morning, so she clears her schedule, makes sure she has a cup of coffee and good breakfast, and gets into the writing. However, engagement is not instant. First, she spends time recalling or reading through what she previously wrote to remember much of her ideas, and then she spends time sketching out what paragraphs to write next. Once she starts producing new text, her immersion still fluctuates; she might switch to other activities due to external interference or self-interruptions. A particular paragraph might be difficult to phrase, and she feels stuck, so she breaks from writing altogether and checks Twitter. Or breakfast wasn't big enough, so she gets hungry and goes for a snack. Perhaps she also has planned interruptions; after an hour she refreshes her coffee, thinks through what to do next, and sits back down.

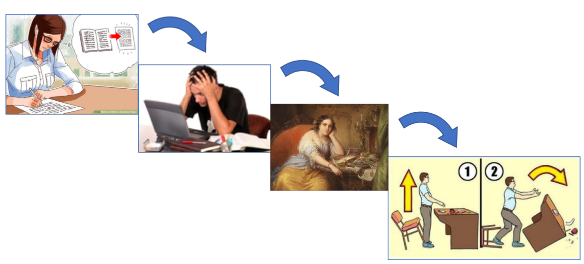

Figure 2.1: Engagement in writing can often move from flow, to split attention, to mind wandering, to disengagement

We engage in a host of different tasks and activities over time — writing a paper, playing a game, eating lunch with friends, or sleeping. We switch from one activity to another, and our absorption in these activities fluctuates. Some of this fluctuation and switching is driven by relatively clear external constraints (e.g., work deadlines) or internal biological drives (e.g., hunger). However, sometimes disengagement is inexplicable; a person might begin work on a paper, but their progress slows, their mind wanders, and finally they decide to switch away to something else. What determines this change in engagement?

Engagement has emerged as a key concept across a broad array of domains, including learning (D'Mello, Dieterle, & Duckworth, 2017), education (Sinatra, Heddy, & Lombardi, 2015; Reschly & Christenson, 2012), the design of games (Boyle, Connolly, Hainey, & Boyle, 2012) and user interfaces (Bouvier, LavouÃ©, & Sehaba, 2014). While it's widely recognized that engagement is not a unitary construct either within or between these domains, there are several general shared characteristics common across all characterizations. Engagement is viewed as a connection between a person and a task, goal, or domain that results in allocation of time and energy towards the foci of the connection. While some uses of *engagement* refer to processes on a timescale of weeks, our interest is in motivational engagement at the level of tasks.

Task engagement refers to a gradient of immersion in a task; from experiencing flow (Csikszentmihalyi & Lefevre, 1989) to partially engaged multitasking (Rosen, Mark Carrier, & Cheever, 2013; Kane, Kwapil, Mcvay, & Myin-germeys, 2007) to full task disengagement. This gradient corresponds to a matching gradient of resource allocation. These task resources include effort, energetic, cognitive, and attentional resources, but also, critically, time resources. *Dis*engagement corresponds a distinct problem of time allocation, when to stop, and we can gain insight into how motivation relates to engagement by separately investigating this scheduling problem.

This paper provides a theory for task engagement based on scheduling of goals. Goal scheduling requires humans to solve a time allocation problem. Many time allocation problems are solved by a priority index assigned to each task or goal, where a worker quits a task when the priority drops below some criteria. Since most features necessary for goal priority are indirect, priority of goals must be inferred. This provides a way of integrating past research in motivation theory based on understanding the *priority cues* (signals that aid priority inference) of different goals that people track. While scheduling and time allocation have appeared in past motivation research (e.g., Simon (1967), Carver and Scheier (1998), Jara-Díaz and Rosales-Salas (2017)), placing it as central in motivational engagement and integrating it with scheduling theory is the goal of this paper.

Human engagement can be viewed as the outcome of an optimal scheduling and resource allocation problem, and our motivational impetus to engage in a task as reflecting the priority of the task's goal. Humans must schedule mutually competing goals that satisfy mutually essential needs; we must decide when to eat, when to sleep, and when to play. To do this people use an array of different signals, like internal hunger or sleep signals or goal progress, as information cues to the priority of different goals.

One goal of this paper is to provide what can be viewed as a *computational* perspective on motivation, and hopefully produce some clarity on how we can connect our subjective experience of motivated engagement with our external activities. While there are a large set of motivational, biological, and cognitive factors that are causally involved, a scheduling perspective can allow us to integrate these together to produce an understanding of engagement via task priority.

2.2 Phenomenon of task engagement

Our above example of the writer at work emphasizes many of the critical components of engagement. Engagement involves a task or activity to be engaged in, in this case writing. While engaged, the activity might subjectively appear easier, and as disengagement occurs it is more difficult to focus. There are nonintentional intrusions into engagement (e.g., boredom or hunger), which can be deliberately worked around through shaping the environment or planning for intrusions. And, importantly, full disengagement results in engagement in another separate task or activity. We could imagine engagement as either a discrete or continuous process (see Figure 2.2), as immersion towards some particular goal and disjoint task switching.

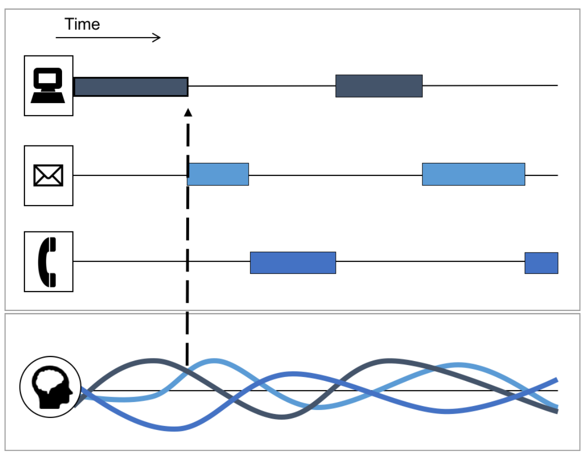

Figure 2.2: (top) A Gantt chart, representing what task a person is engaged with at what time. (bottom) Priority of resource allocation to each task currently in continuous engagement. This motivational engagement drives the more discrete task switching above; when one task achieves highest priority, then a task switch occurs.

Another prototypical example of engagement is video games. While immersed in a Mario game, a player might spend time trying to beat a particularly difficult level that she must attempt while "dying" repeatedly. She might fail multiple times, and get emotionally frustrated, but stay absorbed. Eventually, through effort and skill, she reaches a new level.

Later she might quit half-way through the new level despite not finishing or encountering any particular challenge; she spontaneously quits. The player's engagement, their immersion, effort, and time spent in a game, can fluctuate due to events in the game (e.g., beating a level) and out of the game (e.g., getting tired). The impact of within-game elements are often the focus of game developers, and researchers.

Many people spontaneously engage in game play without external incentives, as attempts to "gamify" education emphasize. *Gamification* tries to capture the relevant aspects of games that promote engagement and apply that to schooling (Dickey & Meier, 2005). The hope is that if we can modify schooling with game features that promote engagement, students might persist with high effort even in spite of task failure, as people often do in games (Hoffman & Nadelson, 2010; Huizenga, Admiraal, & Dam, 2010). This modification generally involves trying to understand the motivational impetus for gaming and translating those motives to educational domains (Ryan, Rigby, & Przybylski, 2006). However, how engagement is defined and measured varies across the literature, in particular between different fields such as pedagogy and game design (Whitton & Moseley, 2014). In education, engagement corresponds more closely to 'time on task', while in game design the focus is on immersion within a task. This can make understanding the motivations underlying engagement more difficult to piece out and translating from one domain to another challenging.

Engagement, in this chapter, refers to the result of resource allocation towards goals. Motivation is often considered the impetus or cause of engagement; motivation towards a task reflects the value associated with that goal (Dickinson & Balleine, n.d.; Colgan, 1989; Hull, 1943; Bolles, 1975). Neither engagement nor motivation are directly (externally) observable; they produce indirect outcomes in terms of what and how people perform different tasks. Motivation and engagement are *latent variables*, i.e., unobservable variables that are not directly measurable via objective methods (see figure 2.4). These latent variables must be inferred in terms of their impact on task choice, time allocation, or task performance, or through psychophysiology and neuroimaging methods. However, they do correspond to particular subjective experiences.

Motivation and engagement are both distinct from subjective enjoyment (Gallistel, 1978; Berridge, 2009; Carver, 2003) and deliberate desire (Mann, Ridder, & Fujita, 2013; Litt, Khan, & Shiv, 2010); however, there is likely an impact at least distally. These dissociations can correspond loosely to different types of utility (Kahneman & Krueger, 2006a); someone's decision utility (utility determining action) might be distinct from their reflective utility (utility revealed retrospectively). These distinctions are important when considering the causal relation between motivation and engagement. Rather than operationalizing engagement via a single measure, we instead focus on how engagement impacts observed behavior via tasks.

What is a task? A focus on laboratory experimentation means that, in cognitive science, tasks are defined based upon explicit goals and instructions an experimenter presents to a participant. Often a task consists of what the participant is rewarded for and what constraints they are under (e.g., (Körding & Wolpert, 2006)). However, beyond the lab, tasks are not as well defined (see e.g., (Lee, Kirlik, & Dainoff, 2013, ch 13)), though there is recent work in artificial intelligence on a formal specification (Thórisson, Bieger, Thorarensen, Siguröardóttir, & Steunebrink, 2016). In our examples, the Mario video game can provide a precise definition of tasks due to the programmatic specification of the game state and the clear success and failures. However, in a more open-world game, or in our writing example, there is far more task ambiguity. Similar difficulties exist in defining an *activity* in time-use studies (Ås, 1978) — how do we extract discrete sequences from a continuous stream of behavior?[1]

In our general case, a task can be considered a distinct segment of behavior that is defined based on a particular goal (Newell & Card, 1985). We provide a clearer definition of goals in section 2.3.1 and a formal expansion of tasks in section 2.4.1, but observationally, tasks can be extracted from the patterns in a person's activity using methodology from computational ethology or activity analysis (Anderson & Perona, 2014; Aggarwal & Ryoo, 2011). While tasks and goals can be hierarchical (Dawkins, 1976; Fentress, 1983; Carver & Scheier, 1998; Simon, 1967; Botvinick, 2008), engagement in daily goals, rather than life goals, are the focus of this paper. Our focus is on tasks concerning more proximate goals at the time-course of a given day, rather than transitions between subtask motor activity or monthly or annual biological and social rhythms. This corresponds to the rational level of timescale as distinguished by (Newell & Card, 1985) and (Anderson, 1990, ch 1), with a focus on distinct or mutually exclusive goals humans can switch between. Therefore our interest in motivated engagement also concerns daily fluctuations in motivation.

2.2.1 Gradient of engagement

While we can describe someone's engagement as all or none, engagement can better be described as a graded phenomenon (see Figure 2.3). Engagement can refer to a broad gradient of different *levels* of engagement, immersion, and absorption in a task or activity (Boyle, Connolly, Hainey, & Boyle, 2012). People can engage in a task to varying degrees, including complete disengagement where people switch from one task to another (Khajah, Roads, Lindsey, Liu, & Mozer, 2016). Psychological flow represents one extreme of "optimal engagement" (Csikszentmihalyi & Lefevre, 1989), where all available cognitive resources are

[1]For this paper, we do not distinguish between an activity or a task. While they have slightly different implications, it is most useful for our case to emphasize the importance of distinct psychological goals as the defining feature of tasks.

allocated and a person is wholly immersed in a particular task. For the video game player, flow involves the external world disappearing, which might only reappear when boredom sets in (Cowley, Charles, Black, & Hickey, 2008). On the other had, off-task mind wandering, as in our writing example, can represent a passive disengagement from a task; while behaviorally the writer is performing the task (e.g., typing on a keyboard), she is allocating minimal effort and cognitive resources without fully switching to a new task (Kane, Kwapil, Mcvay, & Myin-germeys, 2007) [2].

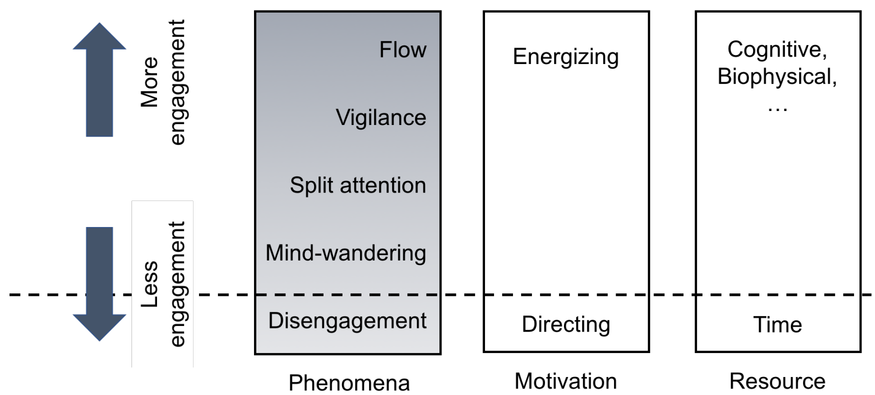

Figure 2.3: The gradient of engagement. Listed are various engagement-related phenomena and where they lie — flow is full engagement, with disengagement at the opposite side. The dotted line represents a bifurcation point in engagement, where someone moves from partial to complete disengagement; they quit the task and are not engaged at all. The motivational distinction between energizing and directing relate to these phenomena and to the distinct resources. Energizing motivation corresponds to allocating more effort, which includes cognitive and biophysical resources. By contrast, directing motivation corresponds to which task is engaged in, which is a question of time allocation (for mutually exclusive tasks). Allocating these resources is a joint problem, as to allocate time requires allocating effort; however, they can appear to have different resulting phenomenon (as seen in the separation of motivation phenomenon into energizing and directing).

This gradient in resource allocation parallels a similar gradient in ecological fear (Mobbs et al., 2015, FEB). Animals have to deal with the cost of predation while still engaging in

[2] Note that mind wandering can also sometimes correspond to a distinct task in some instances.

other activities, and so require a gradient of 'fearful' responses from partial vigilance to imminent defense (Stephens, Brown, & Ydenberg, 2007). Much like engagement, vigilance is operationalized differently in different research contexts (e.g., across species (Lima & Dill, 1990)). In some instances vigilance refers to time spent looking for predators exclusively (as a distinct task), while in others it refers to split attention while performing other tasks (splitting task resources) (Shettleworth, 2010). This "parallel" versus "serial" engagement can also occur via multitasking, and motivation is relevant for either engagement process (Simon, 1994).

Within motivational theory, a distinction is often made between the *directing* aspect of motivation and the *energizing* component (Niv, Joel, & Dayan, 2006). The directing component refers to motivations directed towards distinct tasks (e.g., eating vs sleep), while the energizing refers to the impetus or force applied to a task (e.g., how vigorously to eat).[3] Despite the phenomenological distinction, motivational theories attempt to account for both using an understanding of how motives relate outcomes to their utilities (Berridge, 2004; Niv, Joel, & Dayan, 2006; Keramati & Gutkin, 2014). These different aspects of motivation likely relate to different parts of the gradient of engagement, suggesting that the decision problem of engagement might refer to joint but distinct decisions of both *effort allocation* and *time allocation*.

2.2.2 Engagement and task resources

Engagement is the result of a meta-cognitive process (Cox, Oates, & Perlis, 2011), as the problem of engagement refers to which tasks to engage in and how many resources to allocate, rather than determining the low level actions that solve a given task (see figure 2.5).

[3] Early motivational theories drew heavily from chemical or physical sciences, and considered motivation as a force or energy (Bolles, 1975). This can be seen in (Hull, 1943) explicitly, as his learning equations draw from chemical rate equations to determine rate of behavior.

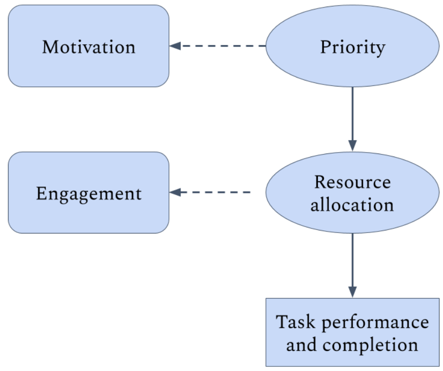

Figure 2.4: Goal priority determines resource allocation (i.e., time allocation) which can determine task performance along with goal satisfaction. However, priority and resource allocation are not directly observable and are instead *latent variables*. The results of the resource allocation, such as task performance and time on task, are observable. Here, we relate the experience of motivation towards a goal as the subjective reflection of the priority that goal has. Similarly, the experience of task engagement is the subjective reflection of various resource allocation (e.g., attentional). Generally, engagement follows motivation. However, in some instances, like externally enforced constraints on time on tasks, engagement and motivation can be disjoint. Similarly, engagement requires resources to allocate; if few resources exist (e.g., due to fatigue), then engagement might be low relative to motivation. Hence we might "desire to do something" separately of actually engagement.

Engagement in a task primarily concerns the consequence of how these resources are allocated across tasks, generally in terms of task completion or performance (see figure 2.4). People who are more engaged in a task allocate more resources, and those less engaged allocate fewer resources, and engaging more resources generally leads to an increase in task completion. However, the types of resources allocated qualitatively change the nature of engagement, which might explain differences in how engagement has been defined across fields.

These different types of resources can correspond to the different aspects of motivation: time resources to directing, and effort to energizing (see figure 2.3). Tasks that are disjoint require directing our time, however once we are performing a task there is an additional choice of how much effort (energy) to expend. Motivational theory generally deals with

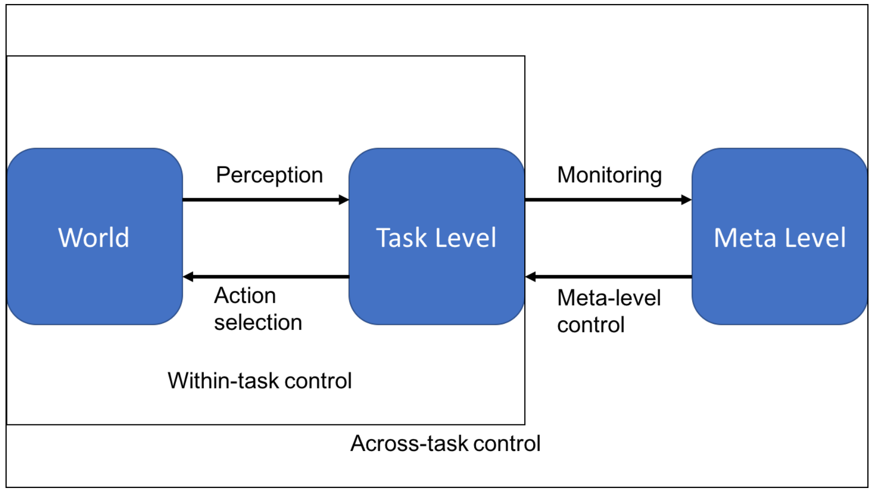

Figure 2.5: A standard control loop, where an agent must interact with the world using actions to solve within-task problems. Outside is a meta-level control that performs across-task control, changing the task controller (e.g., by specifying goals for the task control loop). Engagement is an across-task control problem.

both (Collier & Rovee-Collier, 1983), however the types of tasks studied can lead one to focus on either directing or energizing components (Niv, Joel, & Dayan, 2006). This is because not all tasks benefit equally from both effort and time (see figure 2.6). Some jobs are not substantially impacted by more or less effort beyond a baseline necessary to perform the task. Simple repetitive tasks such as filling out administrative forms or doing laundry certainly require effort, but they may not substantially by improved by *more* effort. This produces a space of different types of resource allocations depending on the type of task, which can change the nature engagement.

Effort, in terms of resources expended, can include biophysical resources like energetic expenditure during overt movement (Shadmehr, Huang, & Ahmed, 2016), but also cognitive and perceptual resources (Sperling & Dosher, 1986). In general, while some mental processes occur automatically (e.g., object recognition), others are limited and must be allocated (e.g., working memory) (Shiffrin & Schneider, 1977). The ability to maintain concen-

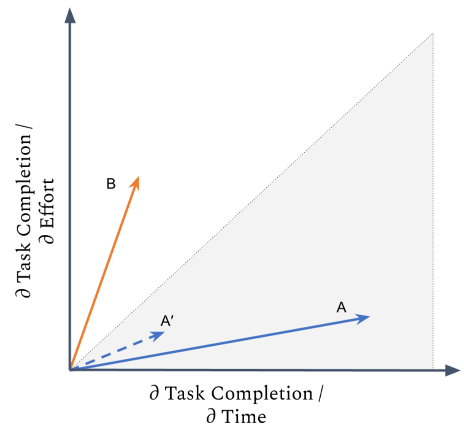

Figure 2.6: Time allocation versus effort allocation in tasks. The horizontal axis is the change in task completion given change in time allocation, while the vertical axis is change in task completion given change in effort allocation to the task. Each vector represents how an individual task's probability of completion is impacted by change in effort and time, or how much the task is benefited by adding more effort or time to it. Some tasks benefit most from increased effort (vector B), while others benefit more from additional time (vector A). For task A, more effort will not substantially increase completion, but time will. The dashed line (A') shows the rate of completion if the task is inactive; there is a small rate of completion even if the task is not being worked on. For this task, there is not a strong gradient of resources that can be allocated at a given time. Instead, this task is either "on" or "off," and just requires scheduling. The grey area represents those tasks that benefit more from time than effort; this is the task region where time allocation matters.

tration or vigilance during a task, often despite task difficulty, can for example be attributed to either energetic expenditure, e.g., neuronal firing (Christie & Schrater, 2015; Matthews et al., 2010), or more abstract cognitive limits, e.g., attentional bandwidth (Matthews et al., 2010; Baker, D'Mello, Rodrigo, & Graesser, 2010; D'Mello & Graesser, 2012). The exact nature of what the resources are that produce the overt feeling of effort at a task is still debated (Shenhav et al., 2017). Traditionally, the concern of allocating these types of mental resources is described in terms of attentional and cognitive control (Sperling & Dosher, 1986).

Cognitive control (i.e., executive function) refers to the set of cognitive process that determines how these limited resources are allocated (Koechlin, Ody, & Kouneiher, 2003).

The relationship between cognitive control and motivation, therefore, investigates how motivation can impact allocation of cognitive resources (Botvinick & Braver, 2015; Braver et al., 2014). Importantly, in most cognitive experimental paradigms, time is held constant and purposefully out of the participant's control (Sperling & Dosher, 1986), making effort allocation the primary theoretical focus. While cognitive control more generally refers to processes involved in goal-driven behavior (Miller & Cohen, 2001), it can be broken down into distinct components such as inhibition, task switching, and working memory processing (Miyake et al., 2000), which themselves might use various resources.

Working memory is perhaps the most well-known limited cognitive resource; people can generally only keep in active thought a limited amount of information (Cowan, 2010; Baddeley, 2000; Ma, Husain, & Bays, 2014). These working memory resources store task-relevant information for a period of time as needed to complete the immediate task — for our writer, they involve both what has been and what needs to be written. However, working memory can also be used for processing information that has nothing to do with the current focal task — "Task Unrelated Thoughts" (TUTs) (Kane, Kwapil, Mcvay, & Myin-germeys, 2007) such as what to make for dinner or whether I need to do laundry. Attentional resources are similar to working memory, in that they refer to a limited bandwidth of information that can be processed, but can be taken up by task irrelevant "distractors".

Task switching with respect to cognitive control refers to how people manipulate and switch mental "task sets" or "schema" (Collins & Koechlin, 2012; Cooper & Shallice, 2006) — a set of cognitive resources, potential solutions, or policies that might solve a given task (Miyake et al., 2000). Task switching as a psychological paradigm concerns the cognitive effects of switching as it relates to these resources (Monsell, 2003).

Research on task switching points to a general set of psychological difficulties and negative impact on performance (Carrier, Rosen, Cheever, & Lim, 2015; Gazzaley & Rosen, 2016). While what the specific resources are is still debated (Koch, Poljac, Müller, & Kiesel, 2018), it is important to emphasize their impact on task behavior and engagement. For our writer this can be clearly seen, as it takes time to actually engage these resources[4]. Again, this is unsurprising as these paradigms are designed purposefully not to study time allocation but rather effort allocation. However, this means there is a theoretical gap, where studies of multitasking often focus on the limitations of mental resources (e.g., (Wang, Irwin, Cooper, & Srivastava, 2015)), rather than addressing the reasons why task switching would occur. This is most prominent in research on self-regulation and inhibition.

Inhibition refers to the disruption of automated processing or prepotent responses by top-down control (Miyake et al., 2000). Hierarchical control in self-regulation theories provides a similar view of inhibition (Hoffman & Nadelson, 2010), in that higher level goal regulation can disrupt currently enacted goal systems (Carver & Scheier, 1998). Certain types of self-regulation is described as costly (Kool, McGuire, Rosen, & Botvinick, 2010) or requiring some sort of willpower (Baumeister & Vohs, 2007) that can be depleted [5]. Some of these resources might be similar in type to biophysical resources, in that an explicit energetic cost might be associated with their use (Christie & Schrater, 2015). For example, Matthews et al. (2010) describe how cerebral blood-flow, which corresponds to energetic "resources" allocated to brain states, also correlates to task vigilance, resulting

[4]Interestingly, people display an inability to anticipate how task switches will reduce their performance on important tasks (Rosen, Mark Carrier, & Cheever, 2013), even though both external and internal distractions are common (Rosen, Mark Carrier, & Cheever, 2013; Judd, 2015; Marulanda-Carter & Jackson, 2012). For example, (Kessler, Shencar, & Meiran, 2009) investigated whether people would switch given the associated performance costs. Importantly while subjects incurred a performance cost due to switching, they still decided to spontaneously switch tasks rather than keep to a single task. While arguably not surprising, (Kessler, Shencar, & Meiran, 2009) claim that this behavior is not predicted based on current theories of cognitive control. "The fact that no one has shown this phenomenon before and that those who studied 'voluntary switching' always took precautions to ensure that participants would switch tasks suggests that this phenomenon has not been predicted" from (Kessler, Shencar, & Meiran, 2009, p127)

[5]While recent failures to replicated ego-depletion bring specifics into question (Hagger et al., 2016), the core idea that energy state can change cognitive function has veracity (Feldman & Barshi, 2007)

performance, and subjective engagement. Overall energy, as indicated by blood-glucose level, has similarly been shown to have strong impact on cognitive performance (Feldman & Barshi, 2007).

While engagement can fluctuate due to allocation of these energetic and cognitive resources, disengagement requires allocating *no* resources to the previously allocated task. The decision to disengage is often viewed as the decision to allocate attentional resources somewhere else (e.g., due to distractions), or as the result of a mechanistic failure. Impulsivity and multitasking, for example, often frame disengagement as a failure of self-regulation (Baumeister & Heatherton, 1996; Baumeister & Vohs, 2007), a failure to properly allocate resources to the given task (Mani, Mullainathan, Shafir, & Zhao, 2013). However, this perspective only considers a focal task and does not consider trade-offs in alternative tasks. The decision to disengage in one task implies engagement in another simply due to the existence of multiple tasks.

2.2.3 Disengagement and multiple goals

Appropriately disengaging from certain goals can be beneficial to individual well being (Wrosch, Scheier, Carver, & Schulz, 2003b), since there are alternatives that can be engaged in. While self-interruptions when multitasking can be detrimental (e.g., (Rosen, Mark Carrier, & Cheever, 2013)), they can also result from trade-offs from multiple goal satisfaction. In a dual-goal setting, for example, people prioritize goals depending on their likelihood of completing both goals (Schmidt & DeShon, 2007; Schmidt & Dolis, 2009). When only one can be completed, time should be allocated to that goal. While there are benefits to engaging deeply in one task, including emotional and motivational benefits (Csikszentmihalyi & Lefevre, 1989), people must switch to satisfy both alternative goals and alternative needs (Wang & Tchernev, 2012), and so people often must engage in these self-interruptions despite being averse to them (Kool, McGuire, Rosen, & Botvinick, 2010). This is not to say that multitasking is not disruptive to goal progress, but that people have more than one goal to complete and so encounter trade-offs.

Recent research in self-regulation has investigated multi-goal pursuit (Neal, Ballard, & Vancouver, 2017; Vancouver, Weinhardt, & Schmidt, 2010; Ballard, Vancouver, & Neal, 2018). This research has had to consider how standard self-regulation processes can be extended to multiple goals, quickly recognizing that there are trade-offs in effort expended across multiple tasks. The existence of multiple goals implies a particular resource trade-off important for disengagement — *time*. The decision task that people might actually be facing can often be different than the one considered by experimenters (Fawcett, McNamara, & Houston, 2012). While single task environments are the standard in most experimental domains, this is the exception rather than the rule in much of life, where multiple tasks constantly vie for our attention. Educational and work domains often treat multitasking as a distraction, given that there is often a focal task of primary importance to either educators or employers. However, both students and employees lead lives outside of work or school, so education and employment tasks may be one task out of many.

These considerations of trade-offs between multiple tasks have been core to issues studied by behavioral ecologists. A central problem animal foragers face is the decision to quit a task for possible alternatives. When should an animal quit foraging from one patch of food and move to another? The irrationality of animals in delayed discounting tasks has been contrasted with the optimality of animals in foraging tasks (Stephens & Anderson, 2001). Delayed-discounting experimental paradigms have been shown to be difficult problems for animals to learn, possibly due to the unnatural structure of the tasks (Blanchard et al., 2015; Carter, Pedersen, & McCullough, 2015)[6]. In foraging, the choices involved take up time

[6]Delayed-discounting tasks are "time pre-allocation" tasks; once time is allocated, you are stuck in either

that cannot be used elsewhere, which produces a trade-off. These constraints can frame the problem differently, suggesting different rational behavior.

The significance here of foraging theory is the focus on time allocation, which appears to be a more natural task for animals than traditional economic choice paradigms. The primary focus of foraging is on understanding how much time animals allocate to tasks (primarily food foraging), though the use of optimal scheduling models (Stephens & Krebs, 1986). In particular, the "patch model" frames foraging as an optimal stopping problem, which results in consideration of a trade-off between foreground tasks and background tasks, where the overall environment and possible tasks that can be engaged in impact the amount of time in a foreground task. For example, possibility of predation or reproduction will impact time of feeding (Stephens, Brown, & Ydenberg, 2007, ch 7). While foraging theory has been extended to human information foraging (Pirolli, 2007) and other more general scenarios (e.g., (Gazzaley & Rosen, 2016)), the importance is not just on the particulars of the "patch model", but rather the ecological validity of the time allocation problem itself and how the choice and stopping problems are distinct.

Here disengagement can possibly be explained with an analogy; people forage across task space much like animals forage across patches. We must make a decision of when to quit a task in the same way animals must decide to quit a patch of food. The number of alternative tasks people can engage in is enormous; people can engage in an unlimited number of possible tasks each day. Consider the game player in the introductory example; while features of the game itself influence how long they play, they have to trade off game-play time with eating, sleeping, working, or other entertainment. Or consider our writer. She explicitly acknowledges this trade-off by clearing her schedule. Many designers in

task with the goal of optimizing bulk reward. Traditionally the delayed-discounting framework is not a time allocation framework, but most experiments functionally require this (as they often require a '"waiting" task) (Stephens, 2008) business and marketing seem implicitly aware of the time trade-off for consumers. Phone application developers attempt to optimize a consumer's time on their own app within an "application environment" (Lewis, 2014; Harris, 2017). In the case of slot machines for gambling, designers will reshape the entire casino (e.g., making exits hard to find and no view of external environment) along with individual slot machines to "push" a customer to stick to the machines (Schüll & Library., 2012).

The choice of engagement as framed above only concerns *which* tasks people engage in, but time allocation also requires us to consider *how long* they engage in each task. While these decisions can be linked, optimal scheduling indicates that they can often be separated out into two problems: a *choice* problem (which task to engage in next) and an *optimal stopping* problem (when to quit the current task) (Gittins, Glazebrook, & Weber, 2011). It is this second problem of time allocation that disengagement must be concerned with, and is also a decision foraging animals must make. A time allocation perspective provides us with a alternative look at disengagement, when combined with our understanding of cognitive and motivational systems. The importance of foraging is that it draws from an ethological view of considering the natural tasks animals engage in when they engage or quit tasks. We must also consider natural engagement, task-switching, and time allocation behavior from people, and look at time as a resource to be budgeted.

2.2.4 Time budgets and human time allocation

Our writer has to trade off their time writing with other tasks, such as eating or entertainment, given a limited time budget. Humans have a wide array of natural tasks they can engage in, making the decision of disengagement one of trading off multiple tasks and choosing to allocate time across them, in addition to allocating other task resources. Animal foraging models analyze time budgets with respect to the animal's ecology (Tinbergen, 1951; Stephens & Krebs, 1986; Colgan, 1989), while human time budgeting is studied across much of the social sciences (Pentland, Harvey, Lawton, & McColl, 1999), including sociology (Ås, 1978), anthropology (Gross, 1984), and economics (Jara-Díaz & Rosales-Salas, 2017). For humans most emphasis is on particular implications of the time allocation for a given field, such as cultural or economic impacts of time use. Empirically these studies often rely on survey data that relies on time-use diaries and other sampling methods, such as the American Time Use Survey as conducted by the American Bureau of Labor Statistics (United States., 2003).

Time use provides a useful window into cultural and historical issues, as time allocation also represents constraints and values that are culturally or environmentally shaped. For instance, the split of our workday into eight hours of work, leisure, and sleep times is a result of historical economics, political policy and culture (Thompson, 1967). Sleep patterns, despite being impacted by natural circadian rhythms (Murray et al., 2009; Webb, Baltazar, Lehman, & Coolen, 2009; Beersma, 1998), are also culturally shaped (Ekirch, 2001). The process of increasing regularity, standardizations, and coordination of our time allocation that occurred in western society is not ubiquitous (Glennie & Thrift, 1996; White, Valk, & Dialmy, 2011). Factory production, agricultural labor, trade and religion all have impacted our time allocation; the work week and calendar celebrations can radically change what tasks we engage in. These particular trade-offs concerning work and labor are explored primarily in economics.

Within economics, time use is studied in terms of how it represents aspects of income, labour, and transportation (Jara-Díaz & Rosales-Salas, 2017). Particular emphasis is on work-leisure trade-offs and intra-household decision making, both as it impacts the larger economy and how it is impacted by sociological identities (e.g., gender). Time budgets are analyzed using decision utility frameworks (e.g., (Becker, 1965; Chiappori & Lewbel, 2015; Heckman, 2015)). These resource-allocation models treat time as a global budget that people spend across tasks with static utilities, dealing with lumped weekly or daily time resources. For example, how much time in a day should you spend on work or travel, given the relative values and costs of each? This contrasts with the dynamic nature of value and engagement that is emphasized in the psychological research (e.g., (Wang & Tchernev, 2012; Adler & Benbunan-Fich, 2013)), or from ecological studies of animal time use.

Natural time allocation behavior in animals is the focus of ethology. Ethological research centers on understanding naturally occurring animal behavior, with emphasis on how instinctual behavior fits into the larger evolutionary and ecological theories (Tinbergen, 1951). Behavior is viewed as another evolutionary adaptation, similar to an animal's physiology, and so requires an understanding of the particular *ecological niche* that an animal is adapted towards (Collier & Rovee-Collier, 1983; Gallistel, 1978). As previously mentioned, foraging theory focuses on time allocation as an optimal stopping problem. Optimal modeling provides a way of connecting the emergent structure of behavior with ecological constraints (Stephens & Krebs, 1986; Colgan, 1989).

The importance of ecological tasks is in providing us with an understanding of what decisions, information, and constraints an animal is naturally going to be concerned with. When considering humans, we must also consider these ecological aspects of time allocation. However, when dealing with modern human behavior, we should be cautious about over-interpreting through some hypothetical ancestral ecology. While history, anthropology, cultural studies, and primatology might be able to provide some insight as to our supposed ecological niche, most suggest the fundamental variety of human time allocation as essential (Gross, 1984). This suggests we should be concerned with more general time budget constraints and structures across tasks.

In recent years, computational ethology has employed machine learning and other computeraided methods to aid in natural behavior description (Andersson, Ramsey, Raemaekers, Viergever, & Pluim, 2012). The primary difficulty in ethology research is the time-labour of detailing behavior. However, given the complexity of "natural behavior," human-coded results were often the only way forward, much like ethnographic methods in anthropology or sociology (Gross, 1984). The rise of computing has allowed these fields to become more quantitative, due to recent advances in activity analysis and modeling. Activity analysis refers to research from more quantitative fields (e.g., computer science or physics), which are interested in either automatically identifying which activity people are engaged in, or forecasting what activities people will engage in (Aggarwal & Ryoo, 2011). In forecasting, traces of behavior are used, such as times a server is pinged, timestamps for emails or tweets, and mobile phone activity (Vzquez et al., 2006; Jo, Pan, & Kaski, 2012; Barabasi, 2005). In these cases tasks are based on the measure used (e.g., emails sent).

Most activity research has found prototypical structure in timing of events that are resulting from human activity (Proekt, Banavar, Maritan, & Pfaff, 2012). These events are non-random in time (i.e., non-Poisson), characterized by bursts of activity followed by heavy tails in times between events (Barabasi, 2005); these features are also found in foraging animals (Jung, Polani, & Stone, 2012). Most models of these events take the form of a time-varying priority queue (Vzquez et al., 2006; Jo, Pan, & Kaski, 2012); people have a short queue of tasks with random completion times, and the priority of each task fluctuates if not completed. These models have been extended to include common periodic structure such as daily wake-sleep rhythms (Kim, Lee, & Kahng, 2013). Otherwise these models don't consider any feature of a task or their value, which point to important commonalities in how people allocate time. We find structure and reliable predictors of human switching and time allocation activity in otherwise unstructured activities such as listening to music (Kapoor, 2014). This points to regularities in the decision problem of disengagement, specifically the idea that we employ a *priority queue* to determine time allocation.

The ways time allocation is investigated vary widely in terms of theory, analysis, and measurement. However there are general aspects that emerge: the importance of a time budget for decisions, the importance of alternative tasks, the dynamic and stochastic nature of interruptions, and how humans have to attend to biological, cognitive, and social goals that themselves differ dramatically. It is worth it for us to consider how humans might be able to *schedule arbitrary goals* to satisfy potential needs, to see if we can gain insight into time allocation. This decision problem is a very difficult one, but regular structure in solutions to scheduling problems points to the use of priority queues as a general solution method to determine when to disengage from a task (Gittins, Glazebrook, & Weber, 2011).

2.3 Modeling the psychology of realistic decisions

A computational modeling perspective can provide us with a framework for engagement as resource allocation by focusing on what the problem is that people solve when they choose to engage or not (Anderson, 1990; Chater & Oaksford, 1999). Human behavior is likely bounded; given the full problem of resource allocation to tasks it is almost certain that humans only approximate the optimal solution. However, considering the optimal time allocation problem provides us with the the landscape of possible theories, and can help us consider if the solution should be of a certain structural form.

Engagement as a phenomenon presents some difficulty for a computational modeling approach. Computational models of human behavior often draw from models of artificial agents, such as using a symbolic logic-based framework or using a Markov Decision Process (MDP) framework (Russell & Norvig, 2009). While symbolic agents have a rich ability to represent goal-space, they often lack motivation or the ability to deal with uncertainty.

By contrast, MDP agents such as those studied in reinforcement learning (RL) require an inflexible reward function on the external environment — that is, a single objective. In order to appropriately represent human engagement behavior, we have to extend these standard approaches.

For humans, task goals form a loose hierarchy that we can navigate – we can flexibly switch between concrete details in action selection and high-level abstract decisions (Koutstaal, 2012). This allows people to construct goals on demand, requiring us to formalize how goals can be specified in a flexible and compositional way (Lake, Ullman, Tenenbaum, & Gershman, 2017). Humans' motives also fundamentally must incorporate an intrinsic component, in that they must be shaped by both internal bodily states and psychological needs that are not directly mapped to the environment (Ryan & Deci, 2000; Berridge, 2004). Importantly, these needs are not necessarily aligned. Humans have exclusive needs that can compete for our time and effort, and it is not a given that they can be stated using the same reward function. This requires us to consider multi-objective optimization, in particular *Pareto optimality*. While these extensions add difficulty, they allow us to more fully represent motivated engagement.

2.3.1 Goals

People engage in tasks to satisfy goals. Our writer wants to complete a paper, and the gamer wants to beat a level. In the broadest sense, goals are the purpose of a task, or the end-point a person wants to achieve. However we want to focus on those goals that are relevant to time allocation. Goals can scale from 'Brush my teeth' to 'Live a good life' or 'Maximize future progeny' (Carver & Scheier, 1998; Botvinick, 2008). Historically goals have a broad use in psychology (Elliot & Fryer, 2008), so here we distinguish between *goal* from an optimization standpoint, as opposed to *goal* from a psychological standpoint (Newell &

Card, 1985; Anderson, 1990).

From an optimization perspective, goals are the objective function that an agent optimizes (Sutton & Barto, 1998). An objective is a function of the agent's state and actions that the agent then acts to maximize (or minimize for costs); the set of actions that maximize the objective function are referred to as the *optimal policy*. From a reinforcement learning perspective, this objective is referred to as *utility* that specifies the goals of an agent (Sutton & Barto, 1998). But we are dealing with the problem of selecting tasks. Each task has its own, local, objective, while the agent has a global objective to optimize. While we might be able to derive all of human behavior from some objective function such as "maximize future evolutionary fitness" (Clark & Mangel, 2000), people construct proximate goals that drive their own behavior (Miller, Galanter, & Pribram, 1960; Carver & Scheier, 1998). Focusing on task engagement requires us to focus on how goals specify a task, and how those goals are represented psychologically.

Psychologically, goals are internal mental representations of a desired endpoint of behavior that impact immediate behavior (Elliot & Fryer, 2008); goals are short-term objectives which are explicitly represented (Gallistel, 1990, Ch 1) (though not necessarily consciously accessible (Aarts & Elliot, 2012, Ch 1)). Goals are often described as set-points of behavior (Carver & Scheier, 1998), in that they produce a feedback loop where behavior is corrected until that set-point is achieved. The set-point itself is some object or state that a person wants to approach (e.g., treasure) or avoid (e.g., monsters) (Elliot, 2006). These types of goals can be most easily seen in motor control (Todorov & Jordan, 2002). Goals in a motor control sense are objective functions that are mapped to physical space. The goal of reaching to a point is represented by having the highest utility (or lowest cost) at that point (Körding & Wolpert, 2006). While this is often combined with some energetic cost, it defines a clear endpoint of the behavior.

However many tasks people engage in have goals that are harder to specify spatially. Many cognitive tasks, such as the Tower of Hanoi, are better specified by logical constraints (Newell & Card, 1985; Anderson, 1993). Symbolic artificial intelligence models often specify goals as a set of constraints on propositions that must be satisfied (Russell & Norvig, 2009). For instance, a goal such as "finish the Mario level without dying" would require atoms $a_1 = finish\ level$ and $a_2 = death$, which are combined into the proposition $g = \{a_1 \wedge \neg a_2\}$. The goal is satisfied if the proposition g evaluates to true (in our case $a_1 = T$ and $a_2 = F$). Traditional AI would search for a plan to execute this goal (Russell & Norvig, 2009), though use of probabilistic logic allows other solutions (Kimmig, Bach, Broecheler, Huang, & Getoor, 2012). For humans, most relevant propositions likely concern states of events (Kurby & Zacks, 2008). This general notion of goals helps us account for many arbitrary tasks humans engage in [7].

For our purposes, goals are internal psychological representations of constraints on events that can be satisfied or not satisfied based on task engagement. Relevant tasks are defined in terms of the goals and current state of the environment, as discussed in Section 2.2. The details of how a goal gets satisfied are of less concern here, only that they can be satisfied by allocating resources towards them. Our focus here is not primarily on how people solve goals, translate goals into actions, or create goals. Rather we are focused on how people prioritize *mutually exclusive* goals; subgoals which are not exclusive (Cooper & Shallice, 2006) can generally be subsumed into a main goal.

Consider our writing example from the beginning. Writing actually represents a hierarchy of goals (Flower & Hayes, 1981). The writer wants to finish an essay, but that itself can be broken down into smaller parts such as "write the introduction paragraph" or "introduce

[7] It is important to mention that constraints can be translated into cost functions formally by using lagrange multipliers, so this discussion more has to do with which is more natural to specify psychological goals.

the problem". These writing goals are generally compound goals — it is not enough to simply finish the introduction but also maintain coherent flow, relate to a specific audience, follow grammatical constraints, and so on[8]. What are relevant for time allocation are those compound goals of the writing task that are small enough to direct immediate behavior, while being mutually exclusive (e.g., the writer generally cannot write both the intro and conclusion simultaneously, but should follow grammar while writing the introduction).

While goals could be created arbitrarily, they have values associated with them, and this impacts whether they are prioritized. Goal representations have multiple components (Moskowitz & Grant, 2009, Ch 1), but include whether they are satisfiable (i.e., whether you can satisfy a goal) and whether they are desirable (i.e., whether the goal is "worth" engaging in itself) (Gollwitzer, 1990). "Can I write this paper?' is distinct from 'Why should I write this paper?" The writing goals above will all impact satisfiability but not desirability. This "desirability" or value is the motivational impetus for a goal (Braver et al., 2014). If this is the case, then how do goals get their value? In other words, what is the criterion or utility that is optimized when allocating time to tasks?

2.3.2 Needs

What determines engagement in a goal? From a time allocation perspective, we are interested in the utilities. To avoid confusion, when we discuss *goals*, we refer to our previous section — goals define the rules of the games people play. The utilities in task scheduling we refer to as *needs* — the reasons why people play games. In a sense, we go up the hierarchy, and ask where goals get their value or utility from. While this could be answered by "other goals" (Carver & Scheier, 1998) or "evolutionary fitness" (Houston & McNamara,

[8]Consider distinct aspects of a grading rubric for writing, or other writing principles (Gopen & Swan, 1990)

2014), and both are in some sense true, the term *needs* allows us to discuss an objective criterion that people might have access to, is appropriately minimal in specification, and does not result in a regress.

Needs here refer to constraints on the internal state of an agent that must be satisfied. Goals that are relevant are therefore those which satisfy needs. While goals refer to external states (e.g., write a paper), needs refer to internal states (e.g., self-actualize). We can specify these needs similar to specifying goals above, as logical propositions over states. This allows us to be minimal in our specification, but still explicit. Need states of a person can be very abstract; your state can refer to your social standing as much as your physical state. Needs are satisfied indirectly by goals; goal satisfaction impacts internal need states by the complex dynamics between our internal and external environments (see figure 2.7).

Needs are relatively stable sets of constraints people must satisfy; how they are determined is less important than that they must be satisfied and therefore determine whether a goal is necessary. Rather than debate over a potential set of foundational needs,[9] the importance is to emphasize the general nature of needs and their impact on goals. We do not refer here to a particular need theory (Bolles, 1975), but rather the idea that there are states of importance for an animal. Needs could instead be referred to as values, but regardless the importance is on detailing how goals receive their importance without having to specify a universal currency for utility (Houston & McNamara, 2014). They are not referred to as drives or motives here because we are discussing *required* states that serve as an objective criterion, not the direct impact on engagement that drives and motives are meant to explain.

As a simple example, the goal of eating your lunch, which is itself a set of constraints, satisfies the need to consume food for homeostasis. While playing a video game satisfies a

[9]For example lists of "foundational needs", see (Reiss, 2004). Interestingly, Maslow explicitly rejects the idea of simply listing needs and instead focuses on what kinds of needs exist (A.H. Maslow, 1943)

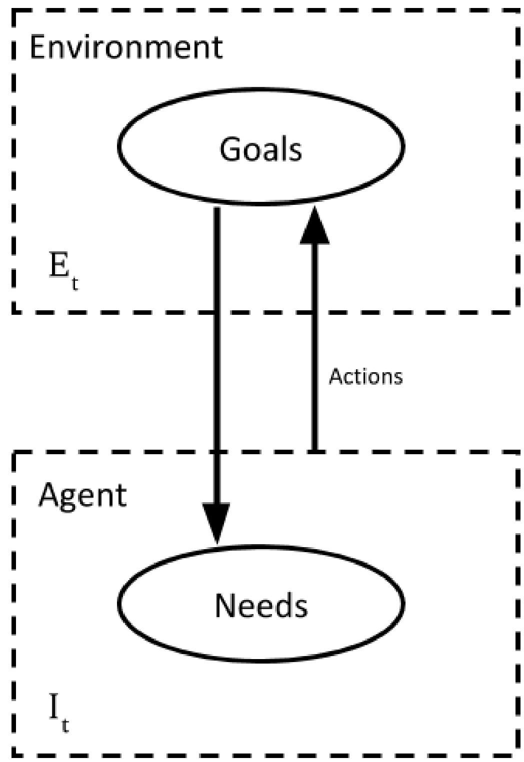

Figure 2.7: Our actions impact goals, and goals impact needs. Needs form a subset from the internal states of an agent at a given time: I_t. While needs are states of an agent, they cannot be directly impacted by our actions. Instead they must be impacted indirectly, via goal satisfaction. The arrows in this diagram refer to causal effect of actions or goal satisfaction. Goals then are subsets of states of the environment E_t. Note that while the *representation* of goals and needs are both necessarily internal to the agent, the constraints those representations refer to are external and internal (respectively).

goal of completing a level, completing the level satisfies the need for mastery. The writing example emphasizes the causal logic of goal and need satisfaction. Our writer might want to finish the essay to satisfy cognitive needs directly (e.g., curiosity or mastery), however publishing the essay might help satisfy career goals (e.g., promotion or tenure, or simply financial stability), which themselves satisfy a complex set of social, cognitive, and homeostatic needs. If she suddenly receives an email telling her she had been fired, her desire to finish the paper (and any imminent engagement) will drop. Breaking the chain by getting fired can make the distal career goals unsatisfiable, which removes the relation between immediate goals and needs (unless writing the paper can satisfy cognitive needs directly). The importance of this causal goal logic is in specifying the relationship between satisfying immediate goals and the satisfaction of needs. People are informationally constrained, in that goal and need satisfaction are uncertain, but humans have access to information that relates to both (Simonov, 1984; Seth, 2013).

Many theories of motivation attempt to explain how a goal's value might change over time, in order to understand temporal changes in behavior independent of task changes (Dickinson & Balleine, 2002; Colgan, 1989; Bolles, 1975). For example, Niv, Joel, and Dayan (2006) formalize motivation as a functional mapping that determines how an agent's internal states impact an agent's reward function, which impacts a goal's value. As internal states change (e.g., hunger increases), this changes the external reward function (e.g., by increasing the value of food). Drive theories mechanistically explain this change in terms of how internal physiological states (e.g., hunger or thirst) impact action selection (Berridge, 2004; Gallistel, 1981). Internal physiological set-points are regulated via a homeostatic feedback mechanism. These homeostatic mechanisms act as basic hardwired controllers that focus on stabilizing and regulating metabolic states. Self-regulation theories of motivation can be seen as generalizations of this set-point framework (Carver & Scheier, 1998), where more abstract social, intrinsic, or identity-based motives specify set-points that then impact immediate goals via progress feedback.

Notably though, while time allocation emerges from these theories (e.g., (Wang & Tchernev, 2012; Vancouver, Weinhardt, & Schmidt, 2010)), they do not specify time as a major *decision variable*. Drives, self-regulation, and motives confuse what is the objective criterion with a separate question; what *information* is used to determine time allocation. Since both needs and goals have their own natural dynamics (due to the dynamics of internal and external environments), this will result in similar time varying allocation, and the question of how this information is transmitted comes back to the front. We will come back to that inference problem in section 2.5, but we will first discuss our specification of the decision problem.

2.3.3 Multi-objective decision making

If an agent has a set of needs that require satisfying, there is no guarantee that a given action will optimize across all needs. Needs now form a multi-objective optimization problem, which has its own distinct set of properties (Sheftel, Shoval, Mayo, & Alon, 2013; Desai, Critch, & Russell, 2018; Roijers, Vamplew, Whiteson, & Dazeley, 2013). For example, imagine our writer only had five hours to split between two different writing tasks, and can make progress in either task based, linearly, on the amount of time spent (i.e., if she spends 1 more hour writing she makes equivalent progress regardless of time spent). Given any time allocation between the two tasks, we can plot the resulting progress in Figure 2.8, where a point on this graph represents the progress achieved after the five hours. Any point above the dashed line is impossible to achieve given our time constraint, and any point below is sub-optimal; we can make more progress by increasing our allocation to either task A or task B. All points on the line then both respect the constraint and use up all available time. If we are only concerned with progress in these tasks, and we have no way of relating progress in these tasks to each other, then all those points are equivalent outcomes. In other words, if we had two utility functions, progress in task A and progress in task B, there is no way to decide which allocation along the line is formally *best*.

In a multi-objective optimization, a decision agent can have more than one objective or utility function. Solutions to the multi-objective are often called *Pareto Optimal* (Sheftel,

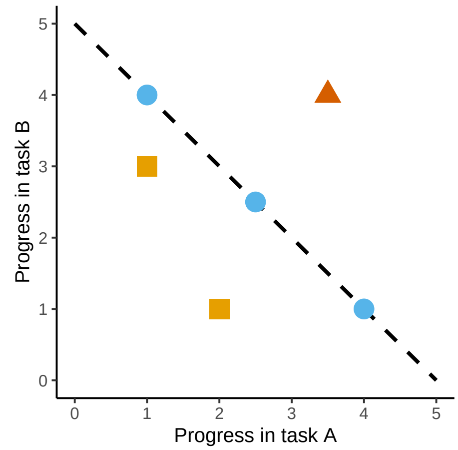

Figure 2.8: Visualization of a simple Pareto Front, plotted in utility space. Each point represents the utility outcome from a high-level action, such as allocating time to two tasks.

Shoval, Mayo, & Alon, 2013). A Pareto optimal decision means there is a trade-off, in that we cannot improve on one utility without reducing the value of another utility, such as the points on the dashed line in figure 2.8. Our writer could improve her progress in one task only by reducing her progress in another. These points form what is called a *Pareto Front*; the set of points which are equivalent in that they are all Pareto optimal. Note that these points represent different policies to take in a decision problem.

One method of solving a multi-objective optimization problem is by relating the different objectives or restating them into one objective. For example, *scalarization* introduces weights β that we can use to add the utilities, e.g., $\mathbf{U} = \beta_1 U_1 + \beta_2 U_2$, where any solution to this equation corresponds to a different point on the Pareto front. In essence, we rewrite our multi-objective problem as a single-objective problem. To continue our example, if our writer could relate progress in each task by the resulting monetary compensation she could get, we could create a single utility function. However, as previously discussed, humans often have incommensurable needs, such as curiosity and financial security. While our writer might be more intrinsically interested in one of the tasks, the other might provide more

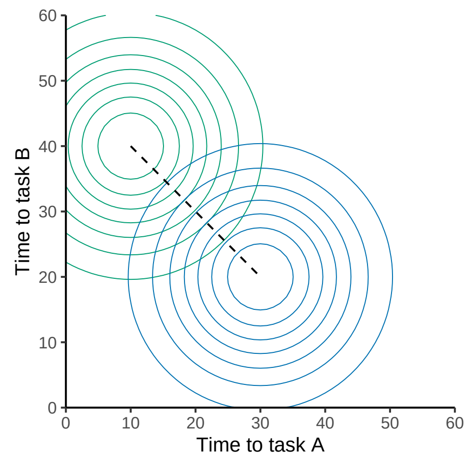

Figure 2.9: Overlay of two different utility functions, given the high-level actions of allocating time to task A and task B. Each contour corresponds to a different utility function, such as resulting curiosity or financial security based on different time allocations. The dashed line is the Pareto front for these utility functions. Note that the Pareto fronts are the maximum across different weighted scalarizations (i.e., different weightings of the two objective functions).

monetary compensation.

Figure 2.9 visualizes the two utility functions as different contours, where each point in the space is a time allocation to the two tasks, with a corresponding value on both utilities (represented as the height via the contour). In this case we may not be able to scale the two functions, for instance if there's no non-arbitrary method of combining needs, resulting in another Pareto front (visualized as the dotted line between the two peaks which form trade offs). This geometry produced by the front again represents all the satisfiable allocation decisions. An important point here is that we have so far treated Pareto in the simplest one-shot decision case. However, Pareto optimality can be computed in the more general multi-step decisions in an MDP (Roijers, Vamplew, Whiteson, & Dazeley, 2013). In those cases, we would instead deal with the expected returns for the utility (i.e., the value function), and these single actions would instead represent choices of policies. Otherwise we may continue with the same analysis, finding the front and set of satisfiable policies.

Multi-objective optimizations represent a more complex set of decision problems, however they also better correspond to the situation most real-world agents are in. Interestingly, many motivational theories attempt to cast all human utility or value into a single unit, such as through evolutionary fitness (Clark & Mangel, 2000). This is unsurprising, as true multi-objective problems are formally very hard to solve. Finding the general Pareto front requires discovering a set of solutions, with no clear method of choosing between them without further assumptions (Roijers, Vamplew, Whiteson, & Dazeley, 2013). Here, we use a multi-objective framework to formalize the resource allocation engagement problem, by allowing our needs to weight the value of goals dynamically. Since we are concerned with need satisfaction, Pareto represents a situation where one cannot satisfy two needs simultaneously. In those situations, our needs induce a trade-off, requiring scheduling.

2.4 Engagement as a scheduling problem

> The *principle of optimal allocation of time and energy* [...] should provide the basis for a general theory of motivation. (Collier & Rovee-Collier, 1983)

What determines engagement in, and disengagement from, a task? Engagement refers to the allocation of coupled but distinct decision variables: task resources (which are themselves varied) and time. This decision problem can also be decomposed into a stopping problem and a choice problem (Gittins, Glazebrook, & Weber, 2011). Disengagement is a bifurcation point in engagement that requires time allocation to be considered — our actions include when to quit a task (i.e., time to tasks). This makes the decision problem of engagement a scheduling problem.

Here we sketch a theory of engagement from a task scheduling perspective, using Pareto Optimality (e.g., Desai, Critch, and Russell (2018)). To understand task scheduling as a decision problem, we first specify what counts as a task by defining goals, and then describe reasons for goal engagement, before discussing the optimal time allocation problem. This allows us to fully specify the relevant states, utilities, and actions in our decision problem. Importantly, we focus on time allocation as the decision to quit a task and, drawing on work from foraging theory and multi-armed bandit problems (Stephens & Krebs, 1986; Gittins, Glazebrook, & Weber, 2011), suggest that the solution to this problem takes the form of a priority queue over tasks. First we expand our understanding of a task by defining goals, and separating them out from the underlying utilities which we call needs.

2.4.1 Decision problem of engagement

One view of engagement is as the result of resource allocation towards a goal. Resource allocation can be viewed as a decision problem — how much of a resource should be allocated to a given task? People are limited in their task resources (e.g., energetic, cognitive, and time), and have a wealth of goals they have to allocate resources towards. Decision models frame behavior as a solution to an optimization problem that maximizes utility of choices. Here the choice is a joint decision of what to do and how many resources to allocate to do it. Adapting decision theory to resource allocation requires specifying the utility of resources in term of the costs and benefits of allocating them, and the ability to forecast future states which affect this value as a function of the resource decision.

Several authors have proposed rational frameworks for making decisions with resource costs (Kool & Botvinick, 2018). While there is much debate about the nature and costs of cognitive resources, it is generally accepted that *time* forms a key resource that may dominate the cost structure. For example, Kurzban, Duckworth, Kable, and Myers (2013) describe how the experience of cognitive cost might result from an opportunity cost from engaging in one task over another. Lieder and Griffiths (2019) formulate a resource-limited single task decision explicitly using decision theory as the choice of a best action plan, which simultaneously maximizes the task utility gained and minimizes an opportunity cost term, construed as a general cost on the time allocated to the task. Their agent's goal is to select a *policy*, π, which maximize its expected utility:

$$\pi^* = argmax_{\pi_T \in \Pi} E(U(\text{results}(\pi_T, s_e))|b, c] - E[\text{cost}(\pi_T, s_e, c)] \tag{2.1}$$

where s_e is the state of the external environment associated with the task, c represents the task context and π represents an extended action plan (i.e., a policy), executed over a horizon T. Note that this policy can include overt actions, cognitive strategy, and resource allocations. The utility U gives the value of achieving the "results" of a task, that is a trajectory of states and outcomes for the task, and the "cost" gives the opportunity costs associated with executing the action plan.

While this structure is quite general for modeling rational resource allocation, adapting the framework to modeling engagement needs to also select between *tasks*, rather than just policies, and specify the rather nebulous opportunity costs. Engagement requires us to consider an agent with multiple, possibly exclusive, goals, while the objective function is specified over multiple needs. If we allocate time T to doing a task, we need to consider what happens over this *episode* both to the task engaged in, and conversely, to the other tasks not engaged in. In so doing, we will show that opportunity costs disappear as an explicit term, and are instead specified as a Pareto competition across the family of alternatives.

Formally, we need to consider three things: 1) need space: the set of dynamic, competing needs which provide a *value* context for engaging in a set of alternatives (without this, an optimization across tasks should produce the one best thing to do); 2) task space, which provides the set of outlets for satisfying our needs; and 3) the episodic returns forecasted for allocating time to these tasks, where we keep track of the *vector* of returns across all the agent's needs. The decision that results is no longer a simple maximization of a single objective – it is a Pareto optimization that jointly selects the task and time allocation in terms of a relative prioritization value.

Under simplifying conditions, this joint decision be separated into task choice, policy choice and time allocation decisions, all of which depend on each other. First we allocate time to tasks and determine for each task the best time investment, and the best policy for that time investment. Normally the policy π selects for actions within a task, however we wish to abstract over tasks. For our approach, we treat the forecasted return from an allocation T to a task j as an important type of value, which we denote by: $K_j(\pi, T) = E[U(\text{results}(\pi_T, s_e))|b, c]$. This is an episodic return for doing a particular task using an action policy π over an episode of length T. While this return is typically a single value, because we explicitly allow our agents to have multiple types of values (those sets of needs), our forecasted returns will always be vectors, with each component representing the value of doing the task with respect to each of a set of objectives. This yields a prospective value for each task under the best allocations for the other tasks (the K tensor). Given these values, the priorities across the tasks can be determined and task choices made. We formalize this problem via a partially observable Markov decision processes (POMDP); a model of sequential decision making where knowledge about the state of the world is uncertain (hence partially observable). A detailed specification of our problem can be found in Appendix 2.7 as a belief MDP, but here we expand on the relevant components.

In our framework, we assume that people have a set of possible goals $g \in \mathcal{G}$. These goals index composite maps over the external environment, specifying which states will satisfy the goal and timing constraints, denoted R_g. Each goal helps satisfy internal reward functions through the coupling between internal need reward functions, indexed by n and external states s_e. (Internal rewards are a set of soft constraint functions, $p(s_i|n)$, and n is a vector of need satisfaction: e.g. $n = [n_1, n_2, ...n_m]$, and each n_i is binary). $n_i = 1$ means the need is satisfied. $\alpha(n, s_e) = \sum_{s_i} p(s_e|s_i)p(s_i|n)$ represents the coupling between internal need satisfaction and external states.

For simplicity, we model the relationship between goals and internal states using a set of weightings α_g on R_g, representing the coupling between the goal and internal states. The result is a *composite utility* function:

$$U(n, s_e, a_t) = \sum_g \alpha_g(n, s_e) R_g(s_e, a_t) \tag{2.2}$$

A composite reward belief-MDP agent is one which tries to optimize the composite reward function $U(n, s_e, a_t)$. Given a slower changing need state, we derive a *conditional* policy which optimizes given a need state, which acts like a context for the agent. We show in Appendix 2.7 we can decompose our global Q-function, which is an average long-term reward given a state-action pair, into goal-dependent functions by weighting via α_g.

Having multiple goals makes this a multiple objective problem, whose solution is termed *Pareto Optimal*. A multiple objective function problem can be expressed in a decision-theoretic form by a weighted combination of objectives via a process called scalarization. Scalarization introduces weights, β_g, that represent the current, *graded* priorities of the different goals g. This multiple objective problem introduces new concepts. Unlike standard decision problems, the weights on the objectives are *free parameters* that need to be optimized. We show that in Appendix 2.7 that our optimal solution across goals takes the form of:

$$\mathcal{K}^*(T_g) = \sum_{t \in T_g} \sum_g \beta_g^* \alpha_g(n, b_t^i) K_g^t \tag{2.3}$$

Where K is the policy-average Q-function for policy g (i.e., policies are matched to goals), and where T_g is a time-scale over which the weights β_g are constant. Equation 2.3 essentially states that we weight our goal-policies K_g^t based on their need-satisfaction α_g and the Pareto weights β_g. When goals are mutually exclusive, that is, when we cannot find policies that allow satisfying multiple goals at once, β_g takes the form of a one-hot vector that essentially selects goals over the time-period T_g. In this instance, the problem becomes one of time allocation, or selection of T_g for each goal g.

The interesting character of the Pareto analysis is that if we change the relevance of a need through α_g, such as by satisfying the internal need state n, this will automatically reprioritize goals by changing the optimal value of β_g, which leads to a *dynamically changing Pareto front*. Reoptimizing the weighting function after changing the need state will lead to task switching if the current goal becomes dominated by the changing Pareto front. Again in our writing task, if our writer suddenly won the lottery she might not need to write if finishing the essay only provided financial stability.

Our objective is to satisfy needs, which we can do by satisfying goals, which we can do by allocating time to goals. This is a dynamic problem in which both the probability of goal satisfiability, and the necessity of need satisfiability, are dynamically fluctuating over time[10]. We can take advantage of some similarities this problem has to the dynamic programming equations of various so-called Bandit problems, a type of time allocation or scheduling problem (Gittins, Glazebrook, & Weber, 2011). Importantly, these scheduling problems have a common solution form, that of a *priority index* over each task to be completed, that

[10]This is similar to a preemptive dynamic policy, with stochastic release dates (Pinedo, 2012, Ch 11). Constraint satisfaction of this kind presents a mixed integer programming problem, unless we soften the constraints by translating unsatisfied constraints into costs or through probabilistic soft logic (Kimmig, Bach, Broecheler, Huang, & Getoor, 2012). Generally speaking, sequencing multiple tasks with uncertain completion times and rewards is a nonlinear programming problem. (Ouelhadj & Petrovic, 2009; Terekhov, Down, & Beck, 2014)

depends only on past information, with time or work allocated to the task with the immediate highest index.

Priority queues as a framework for understanding human task scheduling is not new. Simon (1967) developed a theory of motivation and affect that treats the core problem of motivation as a scheduling problem — allocating time across different tasks based on the task priority, where priority can be impacted by emotion or affect. In order to solve this problem humans maintain a priority queue in memory — each task that should be completed (i.e., goals such as 'eat breakfast' or 'brush teeth') are stored in a queue or ordered list based on a priority attached to each goal. The goal with the highest priority is the goal that is actively controlling behavior. Here we extend this concept formally.

2.4.2 Scheduling as an optimization problem

Since there are mutually exclusive goals that satisfy mutually essential needs, we cannot simply choose one goal over another, as both might be necessary. The only way to solve this is to split our time among distinct goals. Our decision variable then becomes time allocated to goals; we get to choose when to start and when to stop the activation of a goal. Time allocation, or scheduling, is a direct result of having multiple necessary tasks to engage with. We now discuss the implications of equation 2.3. Given research in scheduling and optimal stopping, it is likely that the solution to our time allocation problem takes the form of a priority index.

Scheduling, as an optimization problem, refers to allocating time across multiple different tasks (often called "jobs" in operations research) (Gittins, Glazebrook, & Weber, 2011). This usually involves either sequencing the tasks in the case of tasks with important dependencies (e.g., drying clothes requires washing them), or allocating time to a task in the case of tasks with continuous or uncertain rewards (as in exploration versus exploitation trade-offs). This latter problem is also called an “optimal stopping problem” (Ferguson & Cox, 2012). Most solutions involve optimization methods such as dynamic programming (Bertsekas, Bertsekas, Bertsekas, & Bertsekas, 1995), though particular analytic solutions exist for simpler problems. Insight can be gained by considering these simpler optimal time allocation problems. Actually finding the optimal solution to many practical scheduling problems automatically is difficult due to uncertainty, stochasticity, and the dimension of tasks (Stoop & Wiers, 1996). Practically speaking, simplifying assumptions and heuristics can approximate general solutions in particular domains (Sha et al., 2004).

As previously mentioned, the “patch model” from foraging theory is an optimal stopping problem — when should an animal quit from a patch of food to consume somewhere else? Charnov (1976) derived an optimal solution referred to as the Marginal Value Theorem (MVT) — the optimal amount of time in each patch is when the rate of gains from the patch equal the average overall rate of gains. When the current patch of food becomes worse than the environment average, an agent should quit. Despite simplified assumptions, the MVT accurately describe much of animal foraging behavior (Stephens, Brown, & Ydenberg, 2007), as well as human behavior (Smith, Bettinger, & Bishop, 1983; Pirolli, 2007).

Another type of optimal stopping problem used to investigate human and animal behavior is the multi-armed bandit problem. A player gets to pull from a set of different slot machines (or bandits) in whichever order or as often as they like. Each machine, when pulled, provides a random reward and each machine has their own reward distribution. The solution to the problem of maximizing reward for the player takes the form of computing an index (known as the Gittins index) for each slot machine, and choosing the machine with the highest index at each round (Gittins, Glazebrook, & Weber, 2011). The index is computable from all observed information (a history of rewards from different machines), and is essentially an estimate of the discounted reward rate of each machine, offset by the uncertainty in that estimate. Given this uncertainty, people have to trade off pulling at the best bandit with learning about the reward rates. Bandit problems of this type have been used often to understand the exploration versus exploitation trade-off humans and animals face (Cohen, McClure, & Yu, 2007; Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006), as well as human learning in these environments (Acuña & Schrater, 2010).

Other more general types of bandit problems (e.g. restless bandit, in which the reward distributions change over time) have prompted the development of other indices that might work either as the optimal solution (Gittins, Glazebrook, & Weber, 2011) or as an approximate solution (Sha et al., 2004). A *priority index* solution is common for these more difficult cases, where the index is often based on optimal solutions to simpler models (Pinedo, 2012, Ch 16). The benefit of such an index is that it functions as a sufficienct statistic for the immediate "quality" of a goal. Priority queues have been used as a model for understanding dynamic scheduling systems (Terekhov, Down, & Beck, 2014), and given work modeling human time allocation as a priority queue, this approach appears promising (Jo, Pan, & Kaski, 2012).

We do not pursue a full proof here. However, we note that equation 2.3 satisfies the formal requirements of an index for a type of restless bandit (Gittins, Glazebrook, & Weber, 2011)[11], Instead, we use the value function above to specify a general functional form the

[11]There are various modeling assumptions necessary for any index theorem to be true: 1) rewards are accumulated up to an infinite time horizon, 2) there is constant exponential discounting, 3) there is only one processor (meaning only one task can be worked on at a time) (Gittins, Glazebrook, & Weber, 2011). These are true of our decision problem above, making our policy be a priority index sensible.

priority index should take. The priority index value of a goal can be approximated by:

$$\begin{aligned} u_t^g &= \alpha_g(n, b_t^i)K_g^t(T_g) \\ &= p(b^e|n)K_g^t(T_g) \\ &\approx \mathscr{F}\left(p(n), p(n|g), p(g|T_g)\right) \end{aligned}$$

which we call the goal *urgency*[12], here represented as a forecast of the total need satisfiability of a goal given the time allocated to it (recall that n is a vector of need satisfaction). Note that urgency is a function of a) the need requirement, drive, $p(n)$, b) the probability a goal satisfies the need, goal desirability, $p(n|g)$, and c) the probability a goal can be satisfied, goal feasibility, $p(g|T_g)$. This suggests a specific architecture an agent should have (see figure 2.10). The agent therefore allocates time to goals based upon an estimate of u_t^g, and selects the current goal via a max rule: $g(t) = \max_g u_t^g$.

2.5 Motivational cues for priority inference

If we view priority as a target of computation, we can take an inference perspective on how priority might be inferred. This allows us to reinterpret past research in motivation in a way that's commensurable with a scheduling perspective. Different factors that induce engagement in a human can be thought of as *cues* for inferring the priority of a goal. In other words, many motivational features of a task such as rewards or progress are signals to need or goal satisfaction. Consider what makes our writer quit — a difficult paragraph, hunger, or losing her job. These all change time allocation by impacting the priority of tasks, either directly by changing the priority of writing, or indirectly by impacting the priority

[12] Here we treat $K^*(T_g)$ from equation 2.3 as the unknown optimal value while u_t^g is the approximate or estimate of the priority index.

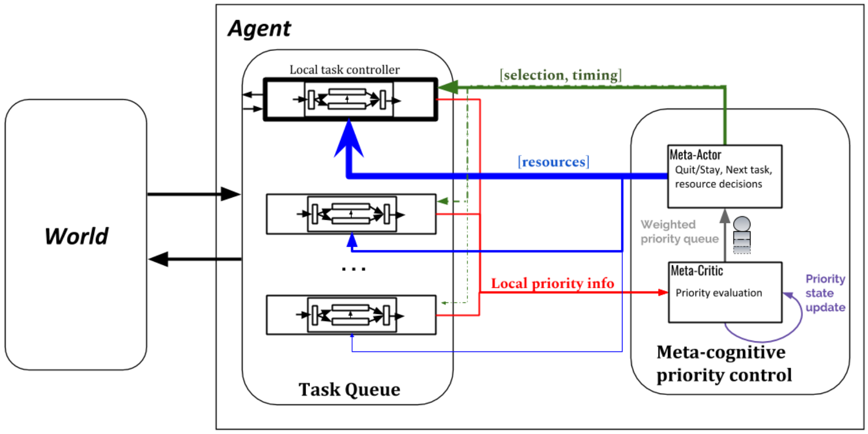

Figure 2.10: Engagement as the result of a metacognitive priority control system. Different local task actors determine which task goals are driving overt actions (within-task control). Outside is a meta-level controller that performs across-task control, assigning resources and activation to each task controller (e.g., by specifying goals for the task control loop). Local task actors send back local priority information from each task to the meta-critic. The meta-critic computes overall priority scores, integrating across longer time scales, while the actor computes priority scores across a salient task set ("queue"). The meta-cognitive system produces action emissions that *must* be coordinated: disengagement, next task selection, resource allocation. Resources are allocated to each local task actor, as well as a "To go" signal, indicating which actor is active (which determines both disengagement and next task selection).

of alternative tasks. This section considers what factors are important for determining the priority of a goal.

2.5.1 Priority inference

In vision research, a cue is a image property that allows you to infer an object or scene property. For example, a cast shadow allows you to determine the depth of an object in a scene (Kersten, Knill, Mamassian, & Bülthoff, 1996). The larger the shadow cast by an object, the farther the object appears from a wall, which can cause the object to appear to move as the shadow changes. Size change is another depth cue; close objects appear larger. Both these cues provide information about the depth of an object, which can be manipulated to produce optical illusions in different contexts. A probabilistic perspective provides a formalism for how these visual cues are combined - as evidence that provides information on object properties (Kersten, Mamassian, & Yuille, 2004). Bayesian inference then allows us to synthesize this information to predict the property from the cues. For example, if you were to combine these shadow and depth cues[13], we would combine the likelihoods $P(c_1, c_2|x) = P(c_1|x)\ P(c_2|x)$ where c_1 and c_2 are the cues and x is the depth we want to estimate. Then we could use Bayes' theorem to produce our estimate of the depth; $P(x|c_1, c_2) \propto P(c_1, c_2|x)\ P(x)$. We can use this same formalism to understand how priority is inferred from cues.

We can consider the causal structure of goals and needs to understand what inference is required for urgency. Consider a student deciding whether to work on an upcoming assignment (see Figure 2.11). The student must infer $p(n), p(n|g), P(g|T_g)$, and marginalize over the different causal paths between the goal (writing an essay) and needs that might be satisfied (here cognitive n_C, social n_S, and physiological n_P needs). We can briefly expand $P(n|g)$ for each goal (those labeled 1, 2, 3). $P(n_C|g_1)$ is direct. However:

$$P(n_S|g_1) = P(n_S|g_4)P(g_4|g_1) + P(n_S|g_3)P(g_3|g_2)P(g_2|g_1)$$

We can similarly expand out for $P(n_P|g_1)$ (not shown). The importance here is the causal structure; as the situation changes, the priority of "Write a good essay" changes based on how that goal relates to needs. If a student's only purpose for writing an essay is to get a good grade in the class, but they find that un-achievable, that will reduce the priority of the essay (unless other social or cognitive needs can be directly tied to essay completion). Similarly, if their financial security is filled by suddenly winning the lottery, the priority

[13] assuming they're independent

of writing can shrink. Note that the actual time allocated will always be a trade off with alternative goals, as time allocation is the result of a priority comparison process.

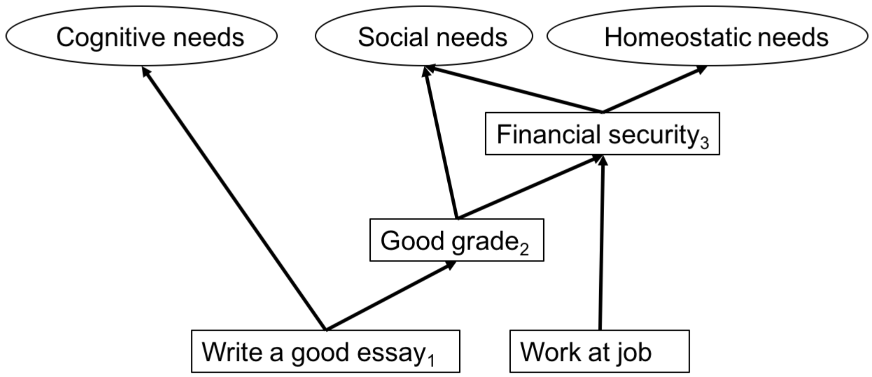

Figure 2.11: The goal of "Write a good essay" for a student, and how satisfying it relates to the satisfaction of other goals (in squares) and needs (ovals). Not all goals and needs are represented in this diagram. The relation between immediate goals and needs might involve a string of logical goal satisfaction. The goal of "Write a good essay" is distally related to financial security, which itself is required for homeostatic needs (among others). For decision making however, this graph can be simplified down into the probabilities relating the immediate goal and needs; how likely will homeostasis, social needs, etc. be satisfied by engagement in writing?

We can expand on the above to deal with temporal inference. Since urgency is a time-varying variable, we can take a Kalman-filter approach to inference, assuming we have a set of urgency-relevant cues or observations o_t. We can assume that our priority is Markovian given each goal's urgency at the last timestep. Given a history of observations, $O_t = \{o_t, o_{t-1}, ..., o_1\}$, we also assume each observation is conditionally independent given the urgency at time t. We can then alternate between a predict and update step. First using a model of urgency transitions, $p(u_t^g|u_{t-1}^g)$, combined with our past inference, we marginalize over the past urgency u_{t-1}^g (predict step):

$$p(u_t^g|O_{t-1}) = \int p(u_t^g|u_{t-1}^g)\, p(u_{t-1}^g|O_{t-1})\, du_{t-1}^g$$

Then update with current observation $o(t)$ using Bayes rule (update step):

$$p(u_t^g|O_t) = \frac{p(o_t|u_t^g)\, p(u_t^g|O_{t-1})}{p(o_t|O_{t-1})}$$

In practice, to implement these equations we would need to specify a particular probability model such that they correspond to the dynamics of goal urgency, for example, satiation and replenishment or periodic dynamics. For now it's important to consider the update equations above in terms of what information should to be computed — the impact of observations o_t on our priority. From an agent-centered perspective, what information should be monitored to appropriately determine priority?

2.5.2 Examples of priority cues

Extrinsic motives

A naive game developer might assume that a main reason people play video games is to garner points (Lewis, 2014). This can often be the first step in gamification and as a framework to understand engagement in casino-style games (Schüll & Library., 2012). Points can function as rewards that provide clear goals and strong feedback, which serve as a basis for reinforcing engagement (McDaniel & Fanfarelli, 2016). However, points only work to engage people in particular situations (Hoffman & Nadelson, 2010) — not all points are treated equally. Points in a game, and reward cues generally, only produce engagement if they are "meaningful," in that the points represent something of motivational relevance for the player. Examples might include social dominance (i.e., having the "high score") or personal skill mastery (i.e., beating your personal best) (Dickey & Meier, 2005). If points are *reliable predictors* of other needs, such as social needs or mastery, then points will be used as cues for priority, and therefore prompt engagement.

The simplest possible cue for priority is an external signal of the availability of a task that satisfies a need — a reward cue. For example, a marshmallow signals the availability of the task "consume a marshmallow," which satisfies a particular homeostatic need. Reward signals are assumed to be privileged observations (Sutton & Barto, 1998), however they only work as a reinforcer if they are reliable cues to satisfying some need state (Rolls, 2009) — reward signals are cues for high probability $P(n|g)$.

These static, external cues are not sufficient to explain engagement, an issue considered by those studying motivational drives (Berridge, 2004). For example, a sudden change in internal state (i.e., salt-deprivation) will cause an animal to re-evaluate the value of a goal (i.e., pulling a lever) by making it high value and prompting engagement (Robinson & Berridge, 2013; Dayan & Berridge, 2014). Preferences for goals, or a goal's reliability as satisfying a need, are dependent on learned context, which can include internal state (Srivastava & Schrater, 2012). When the internal state (and therefore need state) changes, a goal's priority can change. This can produce a seeming re-evaluation of a reward, such as seen in hunger satiation (Keramati & Gutkin, 2014). However other internal needs beyond physiological can also be dynamic.

Intrinsic motives

Our video game player might have suddenly quit part-way through a level because the game became boring — it was too easy or there was nothing to improve at. Video game play is most likely driven by intrinsic motives (Blythe, 2003; Ryan, Rigby, & Przybylski, 2006; Dickey & Meier, 2005; Ducheneaut, 2006). Intrinsic motivation refers to the the set of motives concerning the innate desire to perform a task independent of the task's outcome. In the fields of social psychology and personality psychology, intrinsic rewards are usually contrasted with extrinsic rewards, such as food or money, that are easily modified externally.

In our framework, all needs refer to internal states of an agent, and intrinsic motives refer generally to more abstract cognitive needs, such as those based in curiosity and mastery.

Humans and most animals are intrinsically motivated by a desire to learn (Loewenstein, 1994; Litman, 2005), master and control the environment (Ryan & Deci, 2000; Csikszentmihalyi, 1990) and explore (Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Cohen, McClure, & Yu, 2007), independently of other needs. People play video games if they are at the "right" level of difficulty (Isaksen, Gopstein, & Nealen, 2015; Khajah, Roads, Lindsey, Liu, & Mozer, 2016). Similarly, motivational flow occurs during appropriate challenges that require full attentional resources (Csikszentmihalyi, 1990). Fluctuations in engagement can therefore be driven by both learning and forgetting — boredom (and disengagement) can occur when learning makes the task trivial (Kapoor, 2014). These need dynamics (dynamics in the satisfaction of a need) can influence the priority of a particular goal and produce engagement or disengagement, provided a relevant measure of them can be tracked by the agent.

What many of these intrinsic motives have in common is an information gain. Many computational models that instantiate these psychological theories use a gain in information as a "reward function" for artificial agents (Oudeyer & Kaplan, 2007, November); agents are rewarded for learning more about the environment. Similarly, *empowerment* (Klyubin et al., 2005) formalizes mastery via the mutual information between an agent's actions and perceptions; if an agent can highly impact their perceptions via their actions, they are more empowered. These metrics provide examples of how these intrinsic needs can be computed and could be directly used as cues to update priority. The recent successes of artificial agents at games, combined with these metrics, suggests that video games are a particularly good domain to investigate priority.

Task expectation and uncertainty

Intrinsic motives require an agent to track their own performance or skill. Expectations and progress similarly require humans to track their performance within or across tasks, and the reliability of this feedback impacts engagement (Carver & Scheier, 2002; McDaniel & Fanfarelli, 2016). However while intrinsic cues relate to needs, progress cues relate to goals. Also referred to as *feasibility* (Gollwitzer, 1990), progress indicates probability of goal satisfaction $P(G|w_g)$. Points can also function as signals towards progress, so depending on context signal either goal or need satisfaction. This can be why "word count goals" can provide impetus to writing, as progress feedback is explicit.

People's expectations are incredibly important in understanding time allocation to tasks (Carver & Scheier, 1998; Locke & Latham, 2006). These expectations are informed by environmental factors — framing effects (Tversky & Kahneman, 1986), such as the small area hypothesis, can manipulate progress expectations. For example, being told that you have only 20% more work to do rather than having finished 80% improves engagement (Koo & Fishbach, 2012), despite being descriptions of an identical state. Other within-task factors, such as task progress, are also significant. Task progress provides information about the immediate likelihood of task completion, or goal satisfaction, and so could determine quitting during task engagement. However it is difficult to interpret the impact of progress on engagement, partially due to the fact that it can both increase or decrease time in a task, based on other aspects of expectation (Schmidt & DeShon, 2007).

In a dual task environment the likelihood of completing both tasks changes how progress impacts time allocation (Schmidt & Dolis, 2009; Payne, Duggan, & Neth, 2007). If both tasks can be completed, then time is allocated to the one requiring progress, but if the tasks are mutually exclusive then time is allocated to the one with the most progress. This is highly related to time allocation based on deadlines (Hartonen & Alava, 2013; Jarmolowicz, Hayashi, & Pipkin, 2010), where time on task is decreased when the deadline is far away and increased when it's closer. The importance of the environment in this case is not just in providing information, but in providing *alternative tasks* people can allocate time to. While progress cues provide information about a particular goal's priority, time allocation requires comparing that goal's priority to alternatives.

Environmental priority signals

A critical aspect of foraging theory is the environment's impact on more immediate goals (Stephens, Brown, & Ydenberg, 2007). Animals must contend with the possibility of predation and competing conspecifics, all while simply assessing the overall quality of an environment. Environmental factors are important in both determining the alternative tasks that an agent might choose, and the contextual information an agent has when they make decisions.

Most choices induce certain trade-offs, in particular when there is uncertainty due to limited information (Schmidt, Dall, & Van Gils, 2010). Engaging in one task induces an opportunity cost by not engaging in others, implying that quality of an environment can impact task time allocation by signaling the quality of alternative tasks. This is directly shown in foraging theory's main result, the marginal value theorem, which indicates a trade-off in foreground and background quality on time in task (Stephens & Krebs, 1986). When the environmental quality is high, humans might spend little time in any task (they "skim the fat" off of all tasks). Alternatively in a low quality environment humans might spend more time in each task (they stick to the "oasis in a desert"). While the environment can be literal, as in ecology research, it can also be a more abstract "task environment."

Environmental regularities and personal experience could also shape human decisions by providing us with background knowledge that might bias our decisions (Alexander, Coambs, & Hadaway, 1978; Fawcett et al., 2014). If our perception of the environment's quality is low than the priority of all tasks might be reduced, as in people with depression. This could be an automatic response to a poor environment (Mani, Mullainathan, Shafir, & Zhao, 2013). This was demonstrated experimentally by Kidd, Palmeri, and Aslin (2013), in that children in an unreliable situation rarely waited for a promised reward — satisfying a delayed goal is less likely in an unreliable environment, and so less priority is allocated to it.

The environment also provides signals for many needs that have a periodic tendency. Humans and animals need to regularly engage in goals that satisfy hunger and sleep. Circadian patterns and other daily rhythms are commonly found across behavioral data (Kim, Lee, & Kahng, 2013). These patterns almost certainly derive from the impact of the sun and weather on availability and accessiblity of different goals. If daily rhythms become highly predictible, then it's worth having an internal model of the daily (circadian) cycle and using that to update the priority of tasks. Circadian rhythms can then be synched by so-called *zeitgebers* such as light (Shettleworth, 2010) or even food consumption (Escobar et al., 2011). These rhythms in turn impact how people and animals regulate food (Webb, Baltazar, Lehman, & Coolen, 2009) and sleep (Achermann & Borbely, 2003; Beersma, 1998), and can impact reward motivation more generally (Murray et al., 2009). In fact, models of sleep regulation in humans integrate both a homeostatic and circadian component (Achermann & Borbely, 2003; Beersma, 1998), which might correspond to both the requirement of need satisfaction and the probability the goal can be satisfied respectively.

2.5.3 Emotions, motivation and priority cues

Engagement is not directly driven by emotional states. We previously discussed the dissociation between desire, enjoyment and engagement — between different types of utility (Kahneman & Krueger, 2006b). Research in neuroscience has strongly distinguished between hedonic "liking" versus compulsive "wanting" in animals (Berridge & Robinson, 2003). This dissociation has also been shown in more psychological domains such as alcohol preference (Hobbs, Remington, & Glautier, 2005), intrinsic desires (Litman, 2005), and even political preferences (Winkielman & Berridge, 2003). Enjoyment and desire have somewhat unintuitive impacts on our engagement, as both satisfactory disengagement (Carver, 2003) and jilted engagement (Litt, Khan, & Shiv, 2010) suggest. What might explain this dissociable, yet still important, influence is that priority influences engagement and therefore relevant emotional cues can indirectly impact engagement. While our scheduling is not necessarily based on hedonic priority or some subjectively accessible priority, they are both components of a combined priority schedule.

One view of emotional states is that they are interoceptive signals of internal, often bodily, states (Seth, 2013). However emotions also are impacted by other more external factors such as in environmental fear (Mobbs et al., 2015, FEB). A more general notion is that emotions are informational cues that signal need satisfaction (Simonov, 1984) — emotional states inform us in whether a need is satisfied (where positive valence $\implies$ satisfied and negative valence $\implies$ not satisfied), while the details of the emotional state signal details about which need and how. While the relationship between emotional states and engagement is complex, their importance on motivated behavior makes them an important target for understanding the subjective experience of engagement.

2.5.4 Cognitive control of priority

So far we have treated the priority inference process as implicit and subconcious. We experience priority as motivational impetus, but we can also indirectly change a goal's priority through explicit cognition. While this can occur through proactively reshaping the environment, we also can update by explicitly assessing the probability of goal or need satisfaction. Note that priority cannot be deliberately set directly, as a consequence of priority being an optimal policy. We cannot arbitrarily set our motivation towards some goal. Instead, the inference concerning a priority cue can be updated in various ways. An example already discussed would be in assessing, or reassessing, progress towards some goal.

Cognition allows us to update our inferences by analyzing what might happen, given some model of the world. Decision making can often be solved in either a deliberative or automatic way; either model-based or model -free (Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006). Inferring priority can be accomplished through basic association of past cues, such as learning the association between sugar and a lever. However, forward models are also important in inferring the impact of satisfying certain goals and needs. The inference processes described here can be implicit, involving simple associations or learning and inferring a full causal graph

Consider how the writer can impact their own engagement. Sometimes this is through environmental constraints (e.g., turning off the phone), or satisfying alternative needs that might be distracting (e.g., eating breakfast). Cognitive control of priority can also be achieved by assessing writing progress, or by reminding ourselves of the relation between an immediate goal and a need. Consider again the network of goals and needs in figure 2.11. The distal relationship between writing and financial security might mean writing has a very low priority. However, re-framing "writing a good essay" as a mastery challenge can connect the writing more directly with alternative cognitive needs, boosting the priority. This provides another perspective on the relation between engagement in work or educational domains, where needs are often very distal from the immediate job, and intrinsic motivation (Ryan, Rigby, & Przybylski, 2006).

Another important aspect of cognitive control to point to is the memory resources involved in representing and storing the task queue described in figure 2.10. Representation of future goals involves prospective memory (Einstein & McDaniel, 2005), and the forgetting of goals can be important for understanding why some task switching does not occur. The writer might forget that a section of their paper needs updating, and so never allocate time to it. This seems to be a very overt way in which goals are dropped from the priority cue, and a clear target for how external memory aids can benefit scheduling.

2.6 Discussion

2.6.1 Origins of time allocation problems

Time allocation as a problem emerges due to mutual exclusivity of tasks that can be worked on. Given intelligent beings can reform and modify their tasks, where does mutual exclusivity comes from? From a mathematical perspective, mutual exclusivity emerges from multiple sources of constraints on an agent's control of its internal and external dynamics. Many researchers have previously noted the similarities between various time allocation problems (e.g,. (Addicott, Pearson, Sweitzer, Barack, & Platt, 2017; Mehlhorn et al., 2015; Averbeck, 2015)), and here we relate them.

Foraging-type dynamics: movement constraints

Models based on Markov Decision Processes can be analyzed in terms of constraints on the agent's movements in a state space, encoded in a transition matrix $T(s'|s, a)$. Essentially, the transition dynamics determine when an agent can not be "in two places at once." For example, a major constraint that creates foraging problems is that there are food patches which can't be simultaneously mined (Stephens & Krebs, 1986). Patches represent subsets of states, and foraging problems require a cost of time and energy to move between these patches.

In Markov chain theory, states can be classified into *communication classes* based on whether they transition among each other via the dynamics. If two states can be transitioned between (possibly through other states), then, with a strict cutoff, they are in the same communication class. It is possible to partition states into soft communication classes by thresholding a cost of transitions. These classes form natural subspaces, and tasks which exclusively load onto one subspace can be thought of as *disjoint* from other tasks. This analysis provides a useful characterization of when and why tasks are mutually exclusive. Whenever tasks T_a and T_b involve objectives and controls which lie completely in two disjoint communication classes $s \in C_a$ and $s \in C_b$, the two tasks.

Bandit-type dynamics: information constraints

Another major source of time allocation is due to the communication class structure in belief space. For partially observed environments, the notion of a transition matrix is generalized to transitions in belief space. Belief states represent the agent's understanding of where it is in the space, given the information available. Constraints on how information can be acquired (you can't *view* multiple states at once) impose constraints on belief dynamics.

This is what happens in multi-armed bandit problems, where the agent can only observe the outcomes from the job (process) it's working on, and consequentially can only control its belief state within this job. Information constraints create communication class structure in belief space, which induces a time allocation problem termed *exploration*.

These two types of dynamics form two major ways in which time allocation is studied, foraging due to mutually exclusive states and exploration due to mutually exclusive belief updates (i.e., you cannot learn and do at the same time).

Need dynamics: motivational constraints

What we introduce here is mutual exclusivity in the way internal needs update, when an agent cannot satisfy multiple needs at once. One way of visualizing these constraints is via a linearization of need dynamics, e.g., $n_{t+1} = An_t + g_t^k$, where n_t is the vector need state at time t, A is a transition matrix for the natural need dynamics, and g_t^k is an abstract action, representing the impact engaging in a goal has on the need space (this can correspond to our utility function above, but remapped).

In Figure 2.12 the need space will update based on the choice of time allocation. If time is allocated to either goal 1 or goal 2, progress is only made in one of the dimensions of the need space. This is due to the interaction between goal progress and the coupling between external states and need satisfaction. The simplest way for this to occur is if the two vectors g^1 and g^2 are weakly orthogonal, that is if they load onto at least one independent need (or independent sets of needs). However note that this orthogonality can also depend on the natural need dynamics A; while g^1 and g^2 might overlap, their effective impact for timepoint t is distinct.

What is important about this type of constraint is that it is fundamentally different from the two above. In order to produce non-mutual exclusivity, one would have to modify goals

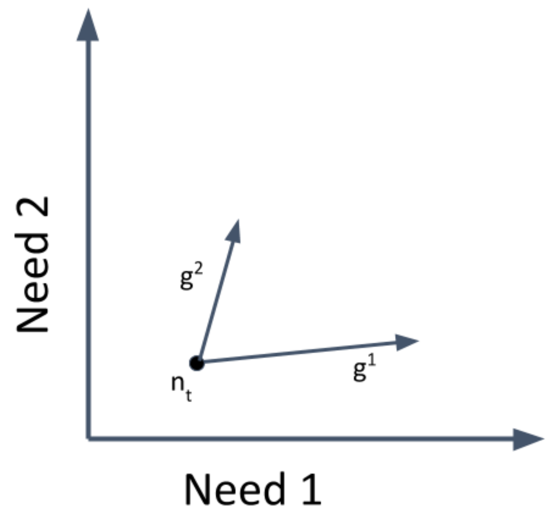

Figure 2.12: Need state as a point in need space, where satisfying either of two different goals results in distinct, approximately orthogonal vectors. Since neither goal can satisfy both needs equally, they are Pareto optimal and produce a trade-off for a single time-point. This means the goals cannot be simply selected between, but rather require an allocation over time.

(not just improve performance towards a set of pre-existing goals). This can help in understanding goal revision, and how fundamental its impact is on time allocation.

2.6.2 Resource allocation as a general problem.

Earlier we made a distinction between different types of resources to allocate, specifically the various cognitive and energetic resources we termed *effort*, and *time*. Our focus in this paper has been on time allocation because we believe it reveals an important structural component of how people engage in tasks. However, understanding how effort is allocated is an equally challenging problem (Shenhav et al., 2017). Time allocation emerges from mutual exclusivity as discussed above; however, there are instances such as in multitasking where tasks are concurrent and require splitting effort (Sperling & Dosher, 1986). Attentional allocation generally has focused on these split efforts partially as an attempt at revealing three things: the underlying structure of attention, whether certain resources can be split, and whether different tasks might share resources.

Even in the case of mutually exclusive tasks, we may need to deal with how effort allocation can impact goal completion. We may have tasks that are highly reliant on both effort and time, as in figure 2.6. In these cases the dynamics of effort allocation can impact the way tasks are scheduled. For instance if cognitive cost depletes energy that requires repletion (Christie & Schrater, 2015), then the sequencing and time allocation of tasks might have long-term impact on the completability of future tasks. We need to rest and recover from some types of effort exertion, and this should be incorporated in future work.

What we have shown as a gradient in engagement is also a continuum between the decision problems of attention allocation and foraging problems; they both represent types of resource allocation with different structural properties due to the types of resources. If priority is changing both types of allocation, as we have stated here, then there can be natural conflations in interpreting the allocation of only one of the resources in isolation. In experimental tasks, researchers often want to independently isolate different psychological and behavioral phenomenon in order to simplify; however, in the case of engagement this might produce improper interpretations. This is especially relevant for fields that extend cognitive tasks to real-world domains, such as in media multitasking (Wang, Irwin, Cooper, & Srivastava, 2015), as the assumptions in the design of cognitive tasks may not hold in real-world situations.

2.6.3 Conclusion

Engagement represents a gradient of distinct resource allocation problems, with the decision to disengage requiring time allocation. Viewing engagement as a scheduling problem can recast motivation to a time allocation perspective, with task priority as a key computation. Many of the standard motivational factors can be considered a type of "cue" to inferring priority. This provides a novel way of integrating seemingly dissimilar motivation research.

This chapter provides a formal sketch of a scheduling theory, which should be expanded upon. A dynamic treatment that restates it as a Markov decision process would provide a stronger grounding. In addition, the independence of other task resources is mostly assumed, but integrating them together would more directly connect this work with recent work in cognitive cost and how that might impact scheduling.

Scheduling and priority provide a way to connect the phenomenology of motivation and engagement with the overt behavior of time allocation. One important take-away should be the importance of an interdisciplinary perspective on the topic of human engagement and time allocation. The question of "what do people do, and why?" is a foundational question of the social and behavioral sciences, as well as the humanities. Taking a fundamentally collaborative view of these otherwise disparate disciplines will promote fundamentally better scholarship. Understanding human engagement should not be a task given to one discipline, especially given the clear ethical importance that the ability to control behavior implies.

Part of this synthesis includes integrating across disparate sources of data. While experimental data is necessary for understanding causes of behavior, observational data can provide us with an understanding of the structure of behavior. While activity and time use data is common within social sciences it is less common in psychology — connecting motivational theories with human time use should be a focus of future work. We also do not discuss the underlying implementation or neurology of priority inference, which should be integrated with our current knowledge of human reward circuitry. Recent work on neuroscience of foraging can be considered in this light. The use of statistics for event-based time series can be a useful tool in this domain (Aalen, Borgan, & Gjessing, 2008), as disengagement and task switching is a punctate event.

Task engagement is a core part of human behavior, and a core difficulty many experience in both subtle and debilitating ways. Scheduling can provide a framework and language for describing these problems, with the hopes of relieving them.

2.7 Appendix

In our framework, we assume that people have a set of possible goals $g \in \mathcal{G}$. Goals index composite maps over the external environment which specify which states will satisfy the goal and timing constraints, denoted R_g. Each goal helps satisfy internal reward functions through the coupling between internal need reward functions, indexed by n and external states s_e. (Internal rewards are a set of soft constraint functions, $p(s_i|n)$, and n is a vector of need satisfaction: e.g. $n = [n_1, n_2, ...n_m]$, and each n_i is binary). $n_i = 1$ means the $\alpha(n, s_e) = \sum_{s_i} p(s_e|s_i)p(s_i|n)$ represents the coupling between internal need satisfaction and external states.

For simplicity, we model the relationship between goals and internal states using a set of weightings α_g on R_g, representing the coupling between the goal and internal states. The result is a *composite reward* function

$$U(n, s_e, a_t) = \sum_g \alpha_g(n, s_e) R_g(s_e, a_t)$$

In the belief MDP framework, the goal maps $R_g(s_e)$ depend on the full belief state, due to possible coupling between internal and external beliefs. Similarly, the full internal-external action vector a_t may affect both types of states due to coupling.

We assume that internal state dynamics are significantly slower than external state dynamics for the purpose of accomplishing simple task-related goals which are the focus of this chapter, like writing, checking email, and eating lunch. Given this assumption, the internal belief state and internal actions will be approximately constant over the time scale of time allocation across tasks, resulting in two time scales t, representing the within task timescale, and τ, representing a coarser scheduling time (e.g., an index) for tasks, and the natural time-scale for internal dynamics. Under this assumption, internal state variables will be approximated as piece-wise constant during epochs of length T_k.

A composite reward belief-MDP agent is one that tries to optimize the composite reward function $U(n, s_e, a_t)$. Given the slow change in need state, we derive a *conditional* policy which optimizes given a need state, which acts like a context for the agent.

Following Braziunas (2003), we specify a belief MDP by the tuple $\langle \mathcal{B}, \mathcal{A}, T^b, R^b \rangle$ where:

- The belief space $\mathcal{B} = \Delta(S)$ given state space S
- $\mathcal{A}$ is the action space.
- $T^B : \mathcal{B} \times \mathcal{A} \mapsto \mathcal{B}$ is the belief transition function where $T^b(b, a, b') = Pr(b'|b, a)$
- $R^B : \mathcal{B} \times \mathcal{A} \mapsto \mathbb{R}$ is the reward function over beliefs: $R^b(b, a) = \sum_{s \in S} b(s) R(s, a)$

Note, however, that we decompose the state space $S = [S_e, S_i]$, that is, the external and internal state space, which produces an identical decomposition of the belief space $\mathcal{B} = [\mathcal{B}_e, \mathcal{B}_i]$.

We want to show how we can decompose a policy over goals.

We now define the Q-function for a need-conditional composite belief-MDP agent. The Q-function is generally defined as:

$$\mathcal{Q}(b_t, a_t) = R^b(b_t, a_t) + \gamma \sum_{o_t \in \mathcal{O}} p(o_t|a_t, b_t) V(b^a_{ot})$$

However, we have a modified reward function above that incorporates need-dependent reward weighting, per goal. We therefore modify the Q-function above to incorporate this new reward function over beliefs:

$$\begin{aligned}\mathcal{Q}(b_t, a_t) &= U(n, b_t, a_t) + \gamma \sum_{o_t \in \mathcal{O}} p(o_t|a_t, b_t) V(b^a_{ot}) \\ &= b_t \sum_g \alpha_g(n, b_t) R_g(b_t, a_t) + \gamma \sum_{o_t \in \mathcal{O}} p(o_t|a_t, b_t) V(b^a_{ot})\end{aligned}$$

We can rewrite this Q-function recursively, by taking averages over the value function on the right of the addition:

$$\mathcal{Q}(b_t, a_t) = b_t \sum_g \alpha_g(n, b_t) R_g(b_t, a_t) + \gamma \bar{\mathcal{Q}}(b_{t+1}, a_{t+1})$$

Where $\bar{\mathcal{Q}}(b_{t+1}, a_{t+1}) = E_{p(o_{t+1}|b_{t+1}, a_{t+1})}\left[\mathcal{Q}(b^{o_{t+1}}_{t+1}, a_t)\right]$, and $b^{o_{t+1}}_{t+1}$ is the belief conditional on next-step future observations. When the observation distributions are concentrated,[14] $\bar{\mathcal{Q}}(b_{t+1}, a_{t+1}) \approx \mathcal{Q}(E_{p(o_{t+1}|b_{t+1}, a_{t+1})}\left[b^{o_{t+1}}_{t+1}\right], a_t)$.

We then proceed with two steps. First, we can specify goal-dependent Q-functions:

$$Q_g(b_t, a_t) = b_t \alpha_g R_g(b_t, a_t) + \alpha_g \left(\gamma \bar{Q}^k_g(b_{t+1}, a_{t+1})\right)$$

Next, If we have a policy library $\pi_k \in$, we can define the per policy return as:

$$\mathcal{Q}^k(b_t, a_t) = b_t \sum_g \alpha_g(n, b_t) R_g(b_t, a_t) + \gamma \sum_{a_{t+1}} \pi_k(a_{t+1}|b_{t+1}) \bar{\mathcal{Q}}^k(b_{t+1}, a_{t+1})$$

Which, for the goal-dependent Q-function, is:

$$Q^k_g(b_t, a_t) = \alpha_g \left(b_t \cdot R_g(b_t, a_t)\right) + \alpha_g \left(\gamma \sum_{a_{t+1}} \pi_k(a_{t+1}|b_{t+1}) \bar{Q}^k_g(b_{t+1}, a_{t+1})\right)$$

[14] Alternatively we can do this by taking equation 42 in Braziunas (2003), and taking a max over the observations and actions.

Theorem 1: We want to show that:

$$\mathcal{Q}^k = \sum_g \alpha_g(n, b^i, a_i) Q_g^k$$

That is, that the global Q-function can be decomposed into goal-dependent Q-functions by weighting via α_g.

Proof: We begin with undoing the global Q-function with a one-step forward look:

$$\begin{aligned}
\mathcal{Q}^k(b_t, a_t) &= b_t \sum_g \alpha_g(n, b_t) R_g(b_t, a_t) + \gamma \sum_{a_{t+1}} \pi_k(a_{t+1}|b_{t+1}) \bar{\mathcal{Q}}^k(b_{t+1}, a_{t+1}) \\
&= b_t \sum_g \alpha_g(n, b_t, a_t) R_g(b_t, a_t) + \gamma \sum_{a_{t+1}} \pi_k(a_{t+1}|b_{t+1}) \\
&\quad \left(b_{t+1} \sum_g \alpha_g(n, b_{t+1}, a_{t+1}) R_g(b_{t+1}, a_{t+1}) \right) \\
&= b_t \sum_g \alpha_g(n, b_t, a_t) R_g(b_t, a_t) + \gamma b_{t+1} \sum_g \alpha_g(n, b_{t+1}, a_{t+1}) \\
&\quad \left(\sum_{a_{t+1}} \pi_k(a_{t+1}|b_{t+1}) R_g(b_{t+1}, a_{t+1}) \right) \\
&= b_t \sum_g \alpha_g(n, b_t) R_g(b_t, a_t) + \gamma b_{t+1} \sum_g \alpha_g(n, b_{t+1}) \bar{R}_g^k(b_{t+1})
\end{aligned}$$

Where $\sum_{a_{t+1}} \pi_k(a_{t+1}|b_{t+1}) R_g(b_{t+1}, a_{t+1}) = \bar{R}_g^k(b_{t+1})$ Then we factorize b using the belief space factorization mentioned above: we can split b into b^i and b^e, where α is based on b^i and Q on b^e (since α requires need states). Here we are assuming two different time scales on the belief space, corresponding to the time scales in state space. This allows us to specify the Q-function just on external beliefs:

$$\mathcal{Q}^k(b_t^e, a_t) = \sum_g \alpha_g(n, b^i) b_t^e R_g(b_t^e, a_t) + \gamma \sum_g \alpha_g(n, b^i) b_e^{t+1} \bar{R}_g^k(b_{t+1}^e)$$
$$= \sum_g \alpha_g(n, b^i) R_g^e(a_t) + \gamma \sum_g \alpha_g(n, b^i) \bar{V}_g^k$$
$$= \sum_g \alpha_g(n, b^i) \left(R_g^e(a_t) + \gamma \bar{V}_g^k \right)$$

Where we absorb the b^e into the R^e.

Therefore:

$$\mathcal{Q}^k = \sum_g \alpha_g(n, b^i, a_i) Q_g^k$$

□

We now focus on how to make a decision given the compound reward function. That is, we demonstrate the optimal decision.

Theorem 2: We want to show that the optimal weighting over the k policies occurs when, if policies are matched to goals, we index into each using a weight-vector β_k^*:

$$\mathcal{K}^*(T_k) = \sum_{t \in T_k} \sum_g \beta_k^* \alpha_g(n, b_t^i) K_{g,k}^t$$

If the policies are mutually exclusive, then the β^* is a one-hot vector that selects goals to pursue, during a time period T_k.

Proof: First, we treat the Q-function as an "instantaneous rate," then we follow the policy associated with k over an allocation T_k. Assume we have a set of policies π_k, and we want to prioritize following policy k by finding a mixture weight β_k. Each policy yields a trajectory distribution for each goal, giving a policy by goal matrix of trajectory distributions.

To more precisely define an expectation over a policy k, given:

$$\mathbb{P}_k(\bar{b}, \bar{a}; \pi_k) = \mathbb{P}(s_1) \prod_{i=1}^{T_k} \mathbb{P}(o_i|s_i)\pi_k(a_i|o_{\leq i}, a_{<i})\mathbb{P}(s_{i+1}|s_i, a_i)$$

where beliefs are $\bar{b} = P(\bar{s}|\bar{o})$. The above definition is a modified version of equation 2 from Desai, Critch, and Russell (2018). Then the expected return from a policy k for the composite reward function is given by:

$$\mathcal{K}_k = \mathbf{E}_{\bar{b}|\pi_k}\left[\mathcal{Q}^k\right] = \int \mathcal{Q}^k(\bar{b}, \bar{a})P_k(\bar{b}, \bar{a}; \pi_k)\mathrm{d}\bar{b}\mathrm{d}\bar{a}$$

plugging in the $\sum_g \alpha_g(n, b^i)Q_g^k$ from Theorem 1 and writing the expectation over times in the allocation T_k explicitly yields:

$$\mathcal{K}_k(T_k) = \sum_{t\in T_k} E_{b_t|\pi_k}\left[\sum_g \alpha_g(n, b^i)Q_g^k(b_t, \pi_k(b_t))\right] = \sum_{t\in T_k}\sum_g \alpha_g(n, b_t^i)K_{g,k}^t$$

where

$$K_{g,k}^t = E_{b_t|\pi_k}[Q_g^k(b_t, b_t)]$$

This is a trajectory average of the Q_g^k function over the trajectories typical from following the policy k. Call these expected value functions $K_{g,k}$.

We can now weight these policies together with a weighting function β_k, where $\sum_k \beta_k = 1$:

$$\begin{aligned}\mathcal{K}(T_k) &= \sum_k \beta_k \sum_{t\in T_k}\sum_g \alpha_g(n, b_t^i)K_{g,k}^t \\ &= \sum_k \sum_{t\in T_k}\sum_g \beta_k\alpha_g(n, b_t^i)K_{g,k}^t\end{aligned}$$

This allows us to specify the optimal weighting β_k^* through Pareto maximization, for a given time selection T_k:

$$\mathcal{K}^* = \sum_k \sum_{t \in T_k} \sum_g \beta_k^* \alpha_g(n, b_t^i) K_{g,k}^t$$

Note that for β_k^* to refer to the global optimal, then $\pi^* \in \Pi$, that is the optimal policy is in our policy library (or available via a mixture of policies from the library). Otherwise this is the optimal with respect to a given policy library.

When policies factorize over goals $g = k$, then this simplifies:

$$\mathcal{K}^*(T_k) = \sum_{t \in T_k} \sum_g \beta_k^* \alpha_g(n, b_t^i) K_{g,k}^t$$

We can further simplify this. If goals are mutually exclusive, that is, if we do not have policies that permit the solution of more than one g at the same time, then the return is maximized by placing all the policy weight over one of the goals, so β_k^* is a one-hot vector selecting goals. If β_k^* is allocated towards one goal g, then $K_{g,k}^t$ represents the value for each goal g, and all unallocated goals have no value. This means that each fixed T_k produces a diagonal matrix of K vectors, one for each goal. Optimizing β_k is then equivalent to maximizing the goal, because the sum does nothing (it's summing a diagonal matrix times a vector, producing a scaled vector).

□

This means that, in situations where goals are mutually exclusive, the optimal solution leads to selecting goals over some time period determined by the scale of internal states T_k, as previously mentioned.

An important note here is that this proof can be extended from discrete to continuous time by initially stating a continuous-time POMDP and using uniformization to enforce the discrete-time steps. Otherwise the proof remains the same.

Given the policy at the lower time scale t, we can now discuss the selection of time allocation over those longer time scales τ, i.e., the problem of scheduling. We initially assumed a particular time constant $T_k < \tau$ where the internal state is unchanging. In the mutually exclusive goal situation, the β_k forms a piece-wise constant solution that is T_k length long, over the longer time-scale.

If we allow selection of T_k, then time should be allocated towards the maximal goal as long as β_k is unchanged (i.e., T_k should be such that $\beta_{k,t}$ is constant for $t \in T_k$). When there exists a shift in the optimal allocation of the β_k^*, that represents a bifurcation in the longer time-scale where $\beta_{k,\tau}$ must transition to $\beta_{k,\tau+1}$, that is, the agent must switch goals. This produces both a time-allocation and goal-selection decision, driven by the internal need states.

Interlude

The meta-cognitive priority control process must monitor background environment and adaptive set decision processes. Decision parameters, such as integration time or attentional resources, should be set in concert with the overall priority of the task. In Chapter 3, we provide a method for measuring the *modulatory processes* responsible for adapting decision effort in response to a distractor task.

Chapter 3

A method for measuring modulatory processes

3.1 Introduction

Decision making requires combining value information with forecasts of potential outcomes (Paulus & Yu, 2012). Human valuation of the environment is complex, incorporating the desire to seek out rewarding items, perform well on particular tasks, and negotiate unforeseen dangers and opportunities. Emotional processes affect decision-making by shaping perceived value via modulatory processes (Phelps, Lempert, & Sokol-Hessner, 2014; Rolls, 2009; Gross & Barrett, 2011). There is abundant evidence that modulatory circuits impact decision making processes through various means (Bogacz & Bogacz, 2007; Basten, Biele, Heekeren, & Fiebach, 2010; Phelps, 2006; Phelps & LeDoux, 2005; Bradley, Sabatinelli, & Lang, 2014), such as by regulating attention (Corbetta & Shulman, 2002), reevaluating tasks and outcomes (Doll, Simon, & Daw, 2012), and changing urgency (Bogacz, Wagenmakers, Forstmann, & Nieuwenhuis, 2010). For example, speed-accuracy trade-offs appear to be modulated by a cortico-striatal network (Bogacz, Wagenmakers, Forstmann, & Nieuwenhuis, 2010; Cavanagh et al., 2011; Berkay, Eser, Sack, Çakmak, & Balcı, 2018) that enables participants to prematurely terminate decisions when speed is needed.

The proven influence of emotional states on decision processes creates difficulties for decision science, as decision processes are typically studied without measuring modulatory emotional processes directly. Moreover, there is not a well developed theory for how and why emotional modulation should impact decision-making. While a given decision task will define much of the decision making process, most models of human decisions involve parameters that can vary depending on factors extrinsic to the task in question. For example, optimal speed-accuracy trade-off depends on the resulting value of different outcomes beyond the decision choice itself (Trimmer, Paul, Mendl, McNamara, & Houston, 2013). Foraging theory suggests that environmental factors should impact immediate decisions, e.g., as the chance of predation increases risky behavior should decrease (Stephens, Brown, & Ydenberg, 2007).

There is increasing evidence that these valuation variables are reflected in a participant's emotional state and mediate decisions via a modulatory process. Unfortunately for researchers, it is difficult to directly measure a participant's emotional state. So, historically, biophysical measures have been used as a proxy ((Eldar, Rutledge, Dolan, & Niv, 2016; Seth, 2013)). However all biophysical measures are indirect and depend on reverse inference (Poldrack, 2006). Two implicit assumptions of many physiological measures of emotional state is that they are 1) independent of each other within a decision context (the *independent measures hypothesis*) and 2) they provide direct access to a given modulatory circuit across decision tasks (the *direct access hypothesis*). Since these are not guaranteed to hold across contexts and measures, we need some alternative method of measuring the latent modulatory process. In this chapter, we aim to provide a means for targeted measurement of the impact of modulatory processes on decision-making.

We propose the use of directed dimensionality reduction, specifically partial least squares, as a method of measuring the impact of modulatory processes on decision making. Directed dimensionality reduction enables the extraction of a decision-relevant latent space from a large set of psychophysiological measures. This bypasses the need to associate one measure with a distinct unobservable process and instead treats our biometric measures as a mixture of decision relevant and irrelevant information. We validate this method using a standard decision paradigm.

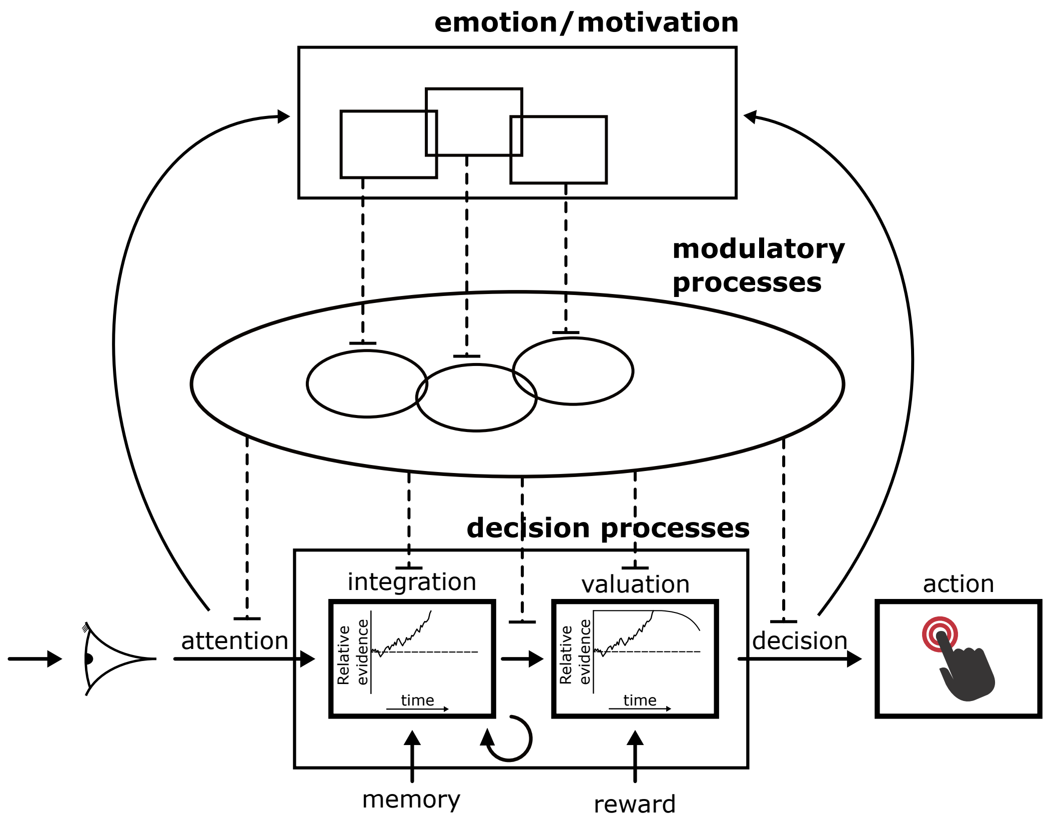

Figure 3.1: Modulatory processes are influenced by emotion/motivation and have coordinated impacts on decision processes. Along the bottom is depicted the drift diffusion model, highlighting the stages where modulatory processes might intercede to impact decisions.

3.1.1 Meta-cognitive control as multiple modulatory circuits

Decision models all share a number of key features (see Figure 3.1). Incoming information is separated into relevant and irrelevant streams, and relevant information is integrated both over time and with memory (Paulus & Yu, 2012; Bogacz & Bogacz, 2007). Decisions are then made through a valuation process that combines integrated information with motivational aspects, such as outcome reward and need satisfaction, until a criterion confidence level is met and an action can be executed. These steps need not be serial, and could involve the integration of information and goals from sources such as input extrinsic to the task demands. Instead, in what follows, we investigate which parts of the decision process are modulated by outside influences.

Decisions are not made in a vacuum. In any decision making environment there are always other potential tasks the agent could engage in. Additionally, the internal state of the decision-maker varies, and how the decision maker is affected by the current external environment fluctuates (Trimmer, Paul, Mendl, McNamara, & Houston, 2013). Emotional states such as arousal, fatigue, anxiety, feeling defensive, or having an intense appetetive drive all affect decision processes. Conversely, previous decisions feed back to influence these emotional states (Seth, 2013). How these emotional and motivational states affect decisions is nuanced; however, there is increasing evidence that decisions are affected through a set of modulatory processes that impact decisions through control loci (Phelps, Lempert, & Sokol-Hessner, 2014).

Coordinating changes in decision-making: The reason for modulatory processes

At a theoretical level, there are advantages to having overarching processes that can coordinate changes in how decision processes work in light of changes in context. To illustrate, consider a typical task such as driving in traffic. From a control perspective, the predictability (or volatility) of other driver behavior is a critical factor that should influence the full range of driving decisions should be made. The presence of an erratic driver on the road decreases the predictive accuracy of internal predictive models. Processes that track the environment are challenged by an increase in the volatility of the environment. This increase in volatility should shorten the integration window over the recent past (Shadmehr & Mussa-Ivaldi, 2012), decrease the accurate predictive horizon, and increase the value of incoming sensory information (i.e., incentivizing an increase in task-related attention). Additionally, it should increase risk of acting by degrading overall certainty, which should increase hesitation (e.g., slowing down to shift away from the erratic driver, engaging in cautious behavior) or even promote switching out of the environment (getting out of traffic entirely). These coordinated changes in information processing and output behavior can be accomplished through modulatory circuits that translate high-level information like environmental volatility into a set of coordinated changes in decision-making circuit function (i.e., change in computations). These high-level cues also covary strongly with emotional state, a relationship we believe is critical to understanding the role of emotion in decision-making.

Some of the ways decisions are controlled are widely recognized. Attentional processes modulate how sensory information is filtered and thus how much information is internally available to the agent at decision time (Yu & Dayan, 2005). The volatility of the environment is monitored by the anterior cingulate cortex (ACC) and other brain areas and influences how information is integrated during the decision period (Botvinick, Cohen, & Carter, 2004) and during the learning of the task (Yu, 2007). Additionally, valuation of options is contingent on the motivational state of the decision-maker (e.g., FeldmanHall, Glimcher, Baker, and Phelps (2016), Tracy et al. (2000)), including the perception of threat or the devaluation of the incentive salience of targeted options due to fatigue or satiety (Lo, Lang, Smith, & Bradley, 2008). Other tasks may intrude, affecting the importance of the decision task relative to other opportunities leading to early quitting or loss of attention (Stephens, Brown, & Ydenberg, 2007, Ch 9). Finally, planning for future demands, such as the resources needed for future activities, can alter the importance and incentive attached to this decision task per se (Christie & Schrater, 2015).

We consider all these decision control factors to be possible loci of modulatory processes that impact decisions by translating emotional, contextual and environmental information into adaptive changes in how and whether a decision is performed.

3.1.2 Instantiating decision process via the drift diffusion model

In order to specify how decisions may be modulated by emotional factors, we capitalize upon a commonly used family of decision making models called drift diffusion models (DDM) (Ratcliff, 1978; Ratcliff, Philiastides, & Sajda, 2009; Busemeyer & Rapoport, 1988; Leite & Ratcliff, 2010) (see Ratcliff, Smith, Brown, and McKoon (2016) for a recent review). Briefly, the DDM illustrates an agent's accumulation of information, quantified as the log likelihood ratio between two competing choices up to the point of decision (see Figure 3.2). Evidence for or against an option noisily accumulates at a certain rate from some starting position. A decision is made once the evidence reaches some threshold (in either the positive or negative direction, which represent each decision). The diffusion model can be derived from sequential sampling theory, specifically the sequential probability ratio test (Ratcliff, Smith, Brown, & McKoon, 2016; Wald, 1945), and reproduces choice probabilities and response time distributions seen in behavior (Ratcliff & McKoon, 2008), overlaid as histograms of correct/incorrect response times in Figure 3.2. The use of this paradigm formalizes a psychological distinction between a person's attentiveness to their environment (the *drift rate* parameter v) and their conservativeness in their decision responses (the *threshold* parameter a). While more general versions of this model includes other parameters, we restrict ourselves to the most common parameter set. Using a parameterized decision model allows us to investigate how an affective/emotional modulatory processes might intervene on specific factors of the decision processes.

These latent decision parameters, drift rate v and evidence threshold a, have reasonably consistent psychological interpretations (Voss, Rothermund, & Voss, 2004) and neural cor-

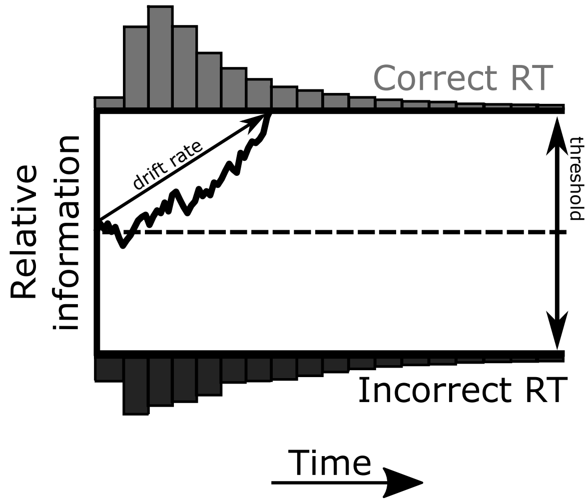

Figure 3.2: Drift Diffusion Model. Two competing choices are represented by top and bottom parallel horizontal lines. Information accumulates noisily over time until a decision is made (the squiggly line crosses either the top or bottom threshold, each representing a competing choice). Example information accumulation is shown, along with resulting response time distributions for both correct and incorrect decisions.

relates (Bogacz & Bogacz, 2007). Various statistical methods allow one to fit the model to observed decision behavior to infer the value of decision parameters. Importantly, we can allow our latent decision parameters to be fit as a regression equation to either experimental manipulations or psychophysical measures. For example, a linear relationship between psychological predictors X and a model parameter θ would be written as

$$\begin{aligned}\theta &= \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \beta_s \\ &= \vec{\beta} X + \beta_s\end{aligned} \tag{3.1}$$

where β_i are regression weights on those measures and β_s is the participant-specific intercept. These decision parameters then impact behavioral response via our understanding of the diffusion model. In particular, it is possible to relate a decision parameter to full distribution through the use of, for example, the *Wiener response time distribution* (Feller, 1968; Luce, 1986):

$$p(t|\theta) = f(t|\theta) + \tau$$

where τ is non-decision time (e.g., motor response time), and $f(t|\theta) = f(t|v, a)$ is the Wiener response time distribution.

For our purposes then, X is a measure of a modulatory process, and $\vec{\beta}_i$ represents its influence on the decision process. For instance, if β is positive, than an increase in the modulatory process increases the parameter. Generally, X is understood to be a proxy of an emotional state or modulatory process. For example, measurements of pupilometry (Cavanagh, Wiecki, Kochar, & Frank, 2014) or functional MRI activity from regions of interests (Frank et al., 2015) have been used as proxy measurements of modulatory processes to test the assumption that these processes impact the decision parameters. The choice of measurement proxy fundamentally limits our ability to investigate the modulatory process.

3.1.3 Emotion and motivation manipulation and measurement

We now discuss how we experimentally manipulate people's physiological state as they perform a decision task, with the goal of collecting relatively large set of brain/body data to facilitate the characterization of emotional state as it influences decision. Emotions reflect many aspects of a person's state, environment, and values, only some of which are relevant for decision making. If we are interested in how emotions impact decisions, we must vary a person's emotional state enough to measure any possible impact the emotional state has on the decision process. We also must measure the set of psychophysiology that is relevant to the emotional manipulation and decision, as these measures provide information about the participant's emotional state.

In our study, participants performed a variation of the random dot motion (RDM) paradigm (Newsome, Britten, & Movshon, 1989) (see figure 3.3). During the task, subject emotional/physiological state was altered via pre-trial auditory stimuli with the goal of capturing how fluctuations in physiological brain/body and emotional states influence decisions.

There is evidence that presentation of loud noises alter both brain/body state and subsequent decisions (e.g., Banis and Lorist (2012), Melamed and Bruhis (1996)). While we use a startle sound to influence a subject's emotional state, the causal relationship between a subject's emotional state and behavior in the task is of primary interest.

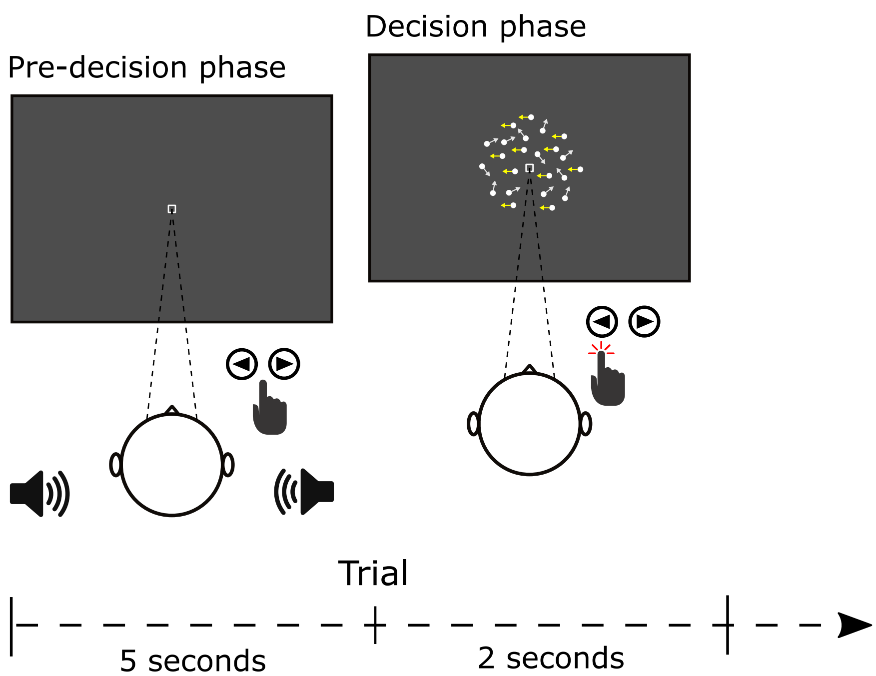

Figure 3.3: Random dot motion decision making task - trial outline. In each trial, we produced a 5 second burst of sound at various levels of intensity (either soft rain, siren, or horn and siren), followed by the RDM stimuli presented for maximum 2.5 seconds (response removes stimuli). After each stimulus presentation we presented 5 seconds of the soft rain sound to bring subjects back to baseline arousal.

We collected a wide set of biometric and psychophysiological measures in order to characterize an individual's brain/body state as it reflects modulatory processes (see Figure 3.4). We collected electroencephalography (EEG), heart rate (HR), galvanic skin response (GSR), facial emotion detection metrics, and pupilometry/eye tracking data throughout the duration of the experiment. These measures were chosen because they have been shown to provide information related to internal brain/body states, such as arousal, (Choe, Blake, &

Lee, 2016; Preuschoff, Hart, & Einhäuser, 2011; FeldmanHall, Glimcher, Baker, & Phelps, 2016; De Vico Fallani et al., 2010), which have been shown to influence decisions (FeldmanHall, Glimcher, Baker, & Phelps, 2016; De Vico Fallani et al., 2010). EEG is a well known method of measuring temporal patterns of neural activity, which may reflect activity in modulatory circuits (Phelps, Lempert, & Sokol-Hessner, 2014; De Vico Fallani et al., 2010). Common biometrics have been used to measure different aspects of peripheral or central arousal, including pupilometry (Preuschoff, Hart, & Einhäuser, 2011; Nassar et al., 2012), HR, and GSR (FeldmanHall, Glimcher, Baker, & Phelps, 2016). Facial expressions are known to express human emotional state (Ekman, 1993), and recent methods of automatic facial detection can do on-line inference of these states from video camera feed.

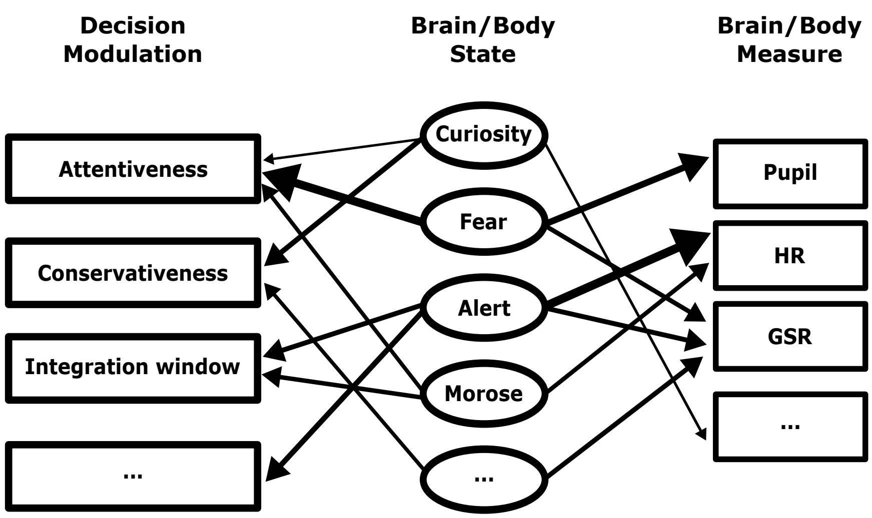

Figure 3.4: Brain/body states influence both decision modulation factors and observable metrics (pupilometry, HR, GSR, etc). Different brain/body states may influence each of these modulatory factors differently, and may be measured to varying degrees as indicated by the weight of the arrows. Feedback between these factors is possible. This figure is for illustrative purposes only and is not intended to make specific claims as to the weights listed.

In order to uncover the impact that modulatory circuits have on the decision process, we regress the DDM parameters against our measures, as in equation 3.1, where biophysical measures are a proxy for modulatory activity, and DDM parameters reflect aspects of the decision process. Past research has characterized how the parameters of the diffusion model, specifically the drift rate and threshold, are dissociable (Voss, Rothermund, & Voss, 2004; Matzke & Wagenmakers, 2009; Milosavljevic, Malmaud, Huth, Koch, & Rangel, 2010) and impacted by various neural circuits (Basten, Biele, Heekeren, & Fiebach, 2010; Bogacz & Bogacz, 2007). Generally speaking, the drift rate appears most modulated by attentional processes or the direct stimuli information itself (Voss, Rothermund, & Voss, 2004; Milosavljevic, Malmaud, Huth, Koch, & Rangel, 2010; Cavanagh, Wiecki, Kochar, & Frank, 2014), while the decision threshold modifies the speed-accuracy trade-off due to time pressure or reward change (Milosavljevic, Malmaud, Huth, Koch, & Rangel, 2010; Voss, Rothermund, & Voss, 2004).

However, there is some research indicating that various manipulations (e.g, emphasizing decision speed versus accuracy) can impact both parameters simultaneously (Rae, Heathcote, Donkin, Averell, & Brown, 2014). This suggests that while these parameters can be, and are formally modeled as, independent, they can also be modulated simultaneously. Many of these studies are hypothesis-driven and were designed to estimate whether the given neural circuit or experimental change had an impact on the relevant decisions. This means that the measures of interest were presumed to contain information about the decision process. We are interested in a different problem: given a set of brain/body data with multiple measures, how do we determine which measures are actually task relevant, and to what degree?

3.1.4 Decision relevant brain/body measures

Decision circuits and brain/body measures are often correlated due to the impact of common modulatory circuits. For instance, general arousal impacts both response thresholds and galvanic skin response (FeldmanHall, Glimcher, Baker, & Phelps, 2016), which produces correlations in their measures. We can therefore gain insight into changes in a decision process through physiological and brain measures.

However, not all of the collected brain/body data will be relevant to a particular decision. To come back to the driving example, a driver's pupils may have dilated in surprise at a sudden shriek of brakes, their heart rate may have increased, and their facial muscles might have contorted at the sound. However, the facial expressions will not have influenced their decisions, whereas the changes in heart rate can (see figure 3.4). Additionally, there will be task irrelevant changes in emotional state from unrelated thoughts (e.g., anger at the other driver). In short, there are brain/body changes that are not part of decision circuits. We want to identify the brain/body measures that directly correlate with decisions, rather than those that are indirectly related as common "downstream" consequences.

In order to address this issue we used biometric measures to predict decisions so that we could get consistent decision-relevant brain/body information. Decision-consistent information can be extracted from noisy biometric data via model comparison, functionally enabling the biometric measures to compete to explain changes in decision behavior. We used the hierarchical drift diffusion model (HDDM) Python toolbox to perform model comparison (Wiecki, Sofer, & Frank, 2013). We performed a series of 3x3 model comparisons, varying each combination of decision parameters (threshold, drift, or both) and decision predictors (stimulus, biometrics, or both). Within each biometric set we were able to determine which parameters/input data best predict decision responses (see 3.5).

We fit all possible models to afford a data-directed determination of whether biometric values provided any information above the auditory stimuli in predicting decisions. Similarly, since a priori either the drift or threshold in the diffusion process could be manipulated by our startle stimuli, we fit models in which either parameter, or both, were treated as a

Decision predictors	Decision parameters: Drift rate (v)	Threshold (a)	Drift rate + Threshold
Stimulus only: 0 - rain, 1 - siren, 2 - siren & horn	Model 1: $v \sim stim_0 + stim_1 + stim_2$ DIC	Model 2: $a \sim stim_0 + stim_1 + stim_2$ DIC	Model 3: $v \sim stim_0 + stim_1 + stim_2$; $a \sim stim_0 + stim_1 + stim_2$ DIC
EEG/biometric data only	Model 4: $v \sim GSR_{pre} + GSR_{post}$ DIC	Model 5: $a \sim GSR_{pre} + GSR_{post}$ DIC	Model 6: $v \sim GSR_{pre} + GSR_{post}$; $a \sim GSR_{pre} + GSR_{post}$ DIC
EEG/biometric data + stim category	Model 7: $v \sim GSR_{pre} \times (stim_0 + stim_1 + stim_2) + GSR_{post} \times (stim_0 + stim_1 + stim_2)$ DIC	Model 8: $a \sim GSR_{pre} \times (stim_0 + stim_1 + stim_2) + GSR_{post} \times (stim_0 + stim_1 + stim_2)$ DIC	Model 9: $v \sim GSR_{pre} \times (stim_0 + stim_1 + stim_2) + GSR_{post} \times (stim_0 + stim_1 + stim_2)$; $a \sim GSR_{pre} \times (stim_0 + stim_1 + stim_2) + GSR_{post} \times (stim_0 + stim_1 + stim_2)$ DIC

Figure 3.5: Nine models were fit for each set of biometric input data, with the GSR set illustrated here as an example. These 3x3 model comparisons assessed which decision predictors (left) and which decision parameters (top) best explain the data. Actual model strings included in each cell. Full model strings for the rest of the biometric input data found in Table 3.15. Models were compared using Deviance Information Criterion (DIC) values. DIC is a common model comparison metric that balances between model fit and complexity (i.e., it penalizes more complex models), and is a hierarchical variant of Bayesian information criterion (BIC). Smaller DIC values mean the model is a better fit for the data. DIC value differences of 5 or more suggest statistical significance. Here (as elsewhere in the paper) we notate a as threshold and v as drift rate in the diffusion model. The stimuli take on three values that are categorically coded in the regression model: calm sound (rain), loud sound (siren), and further loud sound (siren and horn), labeled $stim_0$, $stim_1$, and $stim_2$ respectively. Note that we measured GSR continuously, then computed aggregates over two intervals: before (during auditory presentation) and after visual dot stimulus onset (labeled GSR_{pre} and GSR_{post}).

function of stimulus conditions.

We then repeated this 3x3 model comparison using different sets of brain/body data listed in section 3.1.3 as predictors (heart rate, pupilometry, GSR, facial emotion, and EEG). Within each of these 3x3 model comparisons we used the deviance information criteria (DIC) as a measure of model fit (Wiecki, Sofer, & Frank, 2013) and selected the best-fitting model as our “winning” model for that biometric set. The winning model reveals which parameters/input data best explain the variability in decision responses. We also compared model fit across the biometric sets.

We performed cross-validation to determine which subset of biometric data was most predictive of subject behavior in the task. To do this, we split the data into train/test sets (80/20 percent, respectively). We then fit separate models with different subsets of biometric features, and evaluated the ability of each model to predict held-out data.

Cross validation sets

		Pupil alone	GSR alone	Heart data alone	HPG alone	HPG in PLS	Full PLS
Data inputs	Pupil	✓			✓	✓	✓
	GSR		✓		✓	✓	✓
	Heart data			✓	✓	✓	✓
	EEG						✓
	Facial emotions						✓
Compression	PLS					✓	✓

Figure 3.6: Six different sets of input biometric data were compared via cross-validation. Cross-validation sets are highlighted across the top. Input data into each of these model sets is highlighted along the left. Check marks indicate which data inputs feed into each of the cross-validation sets. Each of the six cross-validation sets has a 3x3 grid of model comparisons as illustrated in Figure 3.5.

3.1.5 Measuring modulatory processes

In order to understand how modulatory processes impact decision-making, we need to characterize and quantify the impact of modulatory processes on behavior. So far we have discussed using proxy measures, testing their impact by using model comparison, and then performing cross-validation to select between predictive measures. For example, heart rate or pupilometry are commonly used as a simple proxy for a state of physiological arousal (Cavanagh, Wiecki, Kochar, & Frank, 2014; Laeng, Sirois, & Gredebäck, 2012), and therefore used to determine how arousal-circuits modulate a particular decision. This is sensible when there is a high correlation between the activity of the measurement and unobserved state, which can be the case within a particular experimental context and paradigm. However, correlation can break down across decision contexts, and requires multiple targeted experiments to piece out how each proxy measure might reflect modulation.

Another tactic is to measure multiple biological features and use them to infer the state of the underlying process. For example, if heart rate and galvanic skin response were highly correlated, that could imply that any variations in their measure is due to a single factor (i.e., a modulatory circuit), as in Figure 3.4. Discovering hidden factors is referred to as factor analysis or *dimensionality reduction*, as one finds the "lower dimensional space" that produces the observed measures.

Biometrics are impacted by other factors besides task-relevant decisions. While heart rate and facial response can both be impacted by a car screech, and both highly correlate, they do not both impact decisions. Standard dimensionality reduction techniques are *nondirected*, and cannot distinguish between decision-relevant correlations. By contrast *directed dimensionality reduction* extracts correlations that are most predictive of a given variable, in our case the relevant decision behavior. By constraining our lower-dimensional subspace to predict task-relevant decisions, we can find the lower-dimensional factors that are task-relevant and impact biometrics. Such underlying latent factors are estimates of a decision-relevant modulatory circuit.

We found that a composite dimensionality-reduced metric is more predictive of behavior than individual biometric predictors. To demonstrate this, we focused on the linear subspace of the cross-covariance that captures key behavioral measures to ensure task-relevance. We employed Partial Least Squares (PLS) to find this linear subspace (Wold, Sjöström, & Eriksson, 2001; Rosipal & Kr, 2006) (see McIntosh and Lobaugh (2004), Krishnan, Williams, McIntosh, and Abdi (2011) for a tutorial and review of PLS in reference to neuroimaging). PLS performs dimensionality reduction and regression simultaneously, and produces latent variables Z (given measures X) that best predict the response Y (see figure 3.7). In our instance, the measures X are the various biometrics and the response Y are decision-relevant behaviors (i.e., response time).

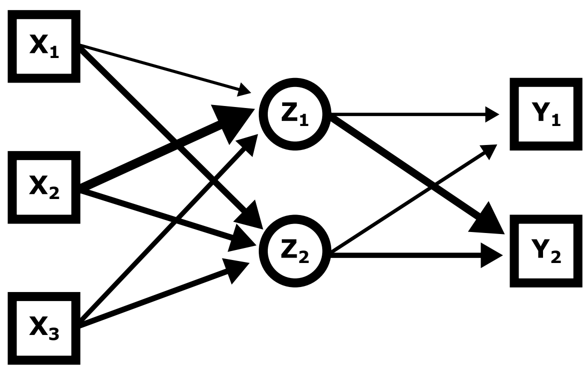

Figure 3.7: PLS demonstration. The X are predictor variables and the Y are responses, while the Z is a latent variable (subspace) meant to explain the relationship between X and Y while accounting for the variance within X and Y. Here the shared subspace Z must respect the connection between X and Y. CCA similarly finds a shared subspace for X and Y, but in that method the width of the arrows are normalized as CCA does not account for the variance in X and Y. PLS, by contrast, cares about the original magnitude of X and Y.

PLS maintains interpretability due to its linearity: the subspace is a linear projection of the measured variables and the latent variables are linearly fit to response. However, many models of decisions imply that the modulatory processes do not have a directly linear impact on decisions (our response variables Y). For instance, in the case of the drift diffusion model, the drift and threshold parameter will change the distribution of response times, but interact in a nonlinear way. While our subspace is predictive of response times, it ignores the decision process captured by the diffusion model. To combine the dimension reduction with the diffusion model, we can simply use the resulting latent variables Z in the diffusion regression equation 3.1:

$$\begin{aligned} \theta &= \vec{\beta} Z \\ \theta &= \vec{\beta} w X \\ \theta &= \omega X \end{aligned} \tag{3.2}$$

Where w is the projection weights found by PLS, such that $Z = wX$, and $\vec{\beta}$ are the weights from the diffusion regression equation 3.1. Also note that due to linearity, we can combine $\vec{\beta} w = \omega$, which simplifies our model and allows us to reinterpret the decision parameters back in the original biometric space. ω captures the impact of the modulatory process on decisions, and allows us to predict how the decision process changes based on relevant biometrics.

This model is fully specified and could be fit as a hierarchical Bayesian model (e.g., fit via MCMC sampling). However this is prohibitive due to the number of parameters needed to both reduce the dimensionality and fit the DDM regression (roughly 10,000). Instead, we take advantage of the above linearity, by first finding candidate latent Z scores and then using those scores as regressors for the DDM model fit. Using only those biometric features that are task relevant (i.e., predict behavior), we can use the nonlinear regression through the diffusion model to estimate impact of those biometrics on latent decision parameters. Since the biometrics used are only those with task relevance, it constrains us to those that are relevant to the modulatory process. This simple linear fit provides a set of candidate subspaces, which can then be compared using standard model comparison methods to identify the subspace most consistent with the data.

Previous research has combined latent variable factor models with decision-process models (Vandekerckhove, 2014; Turner, Rodriguez, Norcia, McClure, & Steyvers, 2016). This is accomplished through simultaneously modeling disparate types of data (e.g., EEG, fMRI, behavior), via the use of latent variables relating these data-sets (Turner, Wang, & Merkle, 2017, July 2016). The present work extends such approaches by using a data-driven method, applied to biometric and neuroimaging data, such that the discovered latent space is constrained to be behaviorally relevant. By contrast, most of these approaches constrained the latent representation of the neural data before relating them to the behavioral data, potentially removing relevant neural information a priori. While this can be appropriate if motivated by theoretical considerations, our approach here requires fewer assumptions to produce a meaningful latent space.

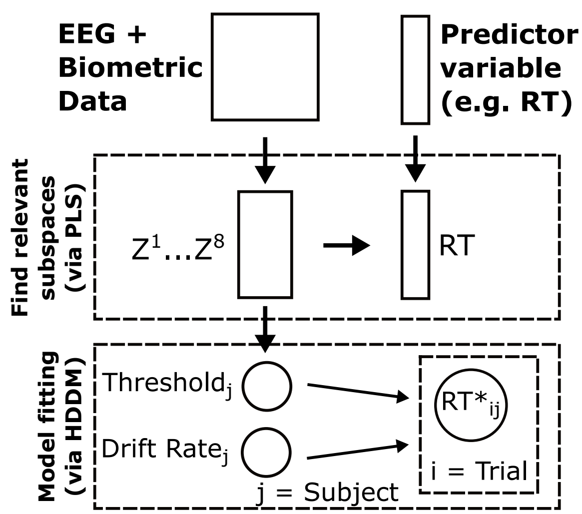

Figure 3.8: Analysis outline. Decision-relevant brain/body data were extracted via targeted dimensionality reduction (PLS). The lower dimensional subset of data, indicated via Z_n, is fed into a hierarchical drift diffusion model (HDDM) comparison, further extracting the components predictive of decision parameters, threshold and drift rate. These were sampled both within and across subjects, hence the “hierarchical” nature of the drift diffusion model parameter estimation.

3.1.6 Validating Methodology for extracting modulatory process

In order to validate our method for extracting modulatory processes using PLS and DDM-based regression on our full set of biometric variables (HR, GSR, pupilometry, facial reaction, and EEG), we compared different sets of biometrics using cross-validation. Each biometric set (e.g., HR vs the full PLS set), was used in the 3x3 model comparison (as explained in section 3.1.4) to determine how that biometric set relates to the decision parameters. Altogether we compared six sets of biometrics: GSR (trial-average), pupilometry, heart rate, all three together, all three through PLS, and the full set including EEG in PLS.

3.2 Results

3.2.1 Which decision parameters and data input predictors are best predictive of decision responses?

We performed model comparison between all data inputs for each model, using the hierarchical drift diffusion model. The HDDM model comparison for model selection are displayed in Figure 3.9. Model selection was based on choosing smallest DIC score within each biometric input set (e.g., set of regressors). Note that for all data input, the stimuli-only models (top row in Figure 3.5) have nearly identical DIC scores (hence are overlapping in the figure). This is trivially true since these models do not use the different data input types (as they are stimulus-only models), and it was performed intentionally to assess variability between model runs with identical inputs. Models 1, 2 and 3 were each run 6 times and had DIC means of 2767.71, 2756.89, and 2787.84, with standard deviations of 0.366, 0.585, and 0.433, respectively.

Based on DIC score, the best-performing model used the full set of biometric data, reduced through PLS, using both drift-rate parameters. However, DIC scores are not comparable when different input data are used. In order to account for possible inaccuracies in model comparison across data input, we performed cross-validation as an additional measure of model quality.

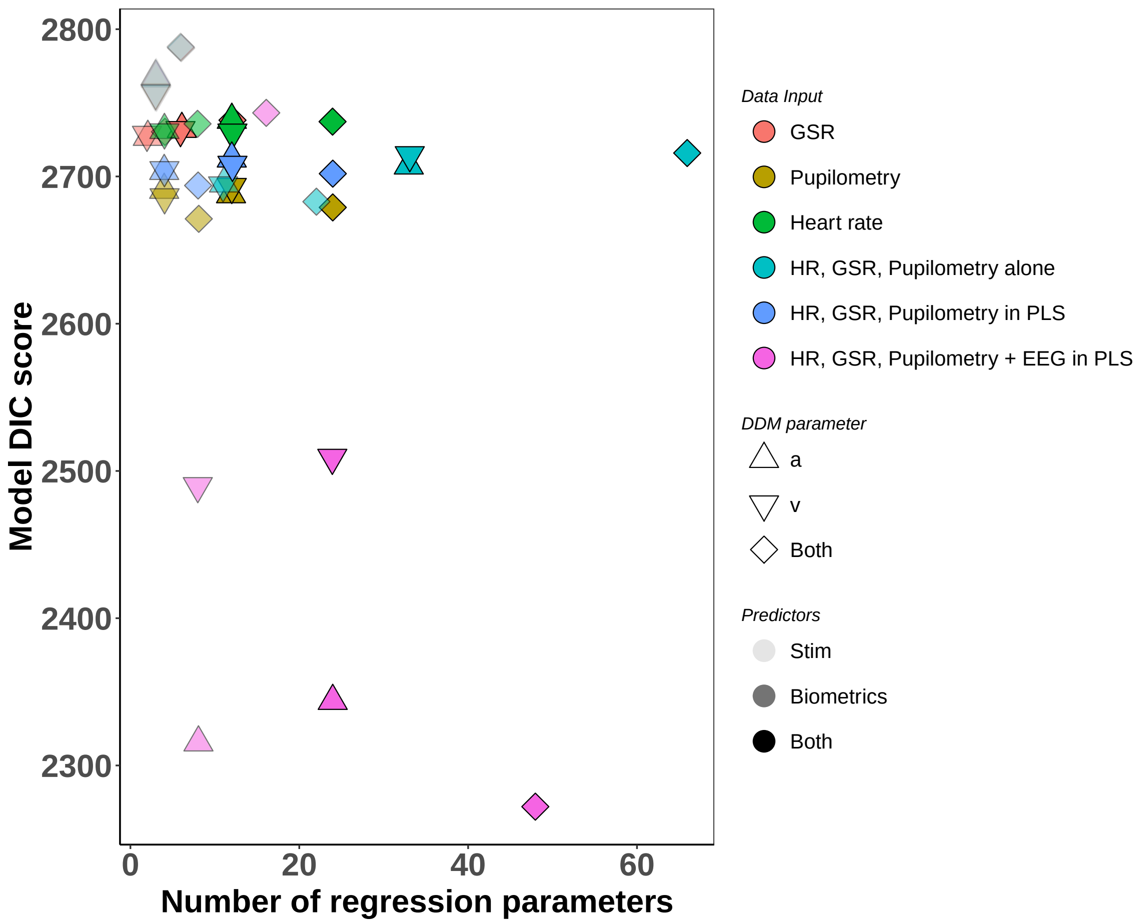

Figure 3.9: Plotted DIC scores of each fit HDDM model. Each point represents the DIC score for a given model, plotted based on the number of input parameters into the regression models (note that this only considers regressors of the *a* or *v* parameters, not hierarchical parameters of the HDDM, which are equivalent across all). Shape indicates which drift diffusion parameter is regressed against, and shading indicates whether the biometrics where used as inputs into the regression (versus the trial stimuli or both). Each color indicates which biometrics were used for the regression. A lower DIC value indicates a better model fit.

3.2.2 Which set of brain/body data is best predictive of decisions?

As HDDM model comparison can only provide information as to which model within a set is most predictive of decision responses, we performed cross-validation to enable the comparison across each data input (e.g., compare GSR and HR performance). The main results of the cross-validation can be seen in Figure 3.10. As described in the methods, we computed the likelihood of observed reaction time given regressors in the test datasets over the fit model parameters (see supplemental eq 3.4 for averaging), for the 31 cross-validated 80:20 splits for each biometric input set. We computed performance relative to average performance for a specific test data set (since all biometric input models were compared on the same train:test splits). Performances relative to average for each biometric input were as follows: Full PLS mean 1.09 (0.023 stdev), HPG PLS mean 0.993 (0.006 stdev), HPG alone mean 0.994 (0.007 stdev), Pupilometry mean 0.998 (0.008 stdev), GSR mean 0.983 (0.006 stdev), and HR mean 0.945 (0.016 stdev) (see Figure 3.10). Note that PLS full has the highest performance, consistant with our model comparison using the DIC metric above.

3.2.3 Quality validation for the discovered latent modulatory process

Dimensionality reduction was utilized in some of the brain/body data subsets analyzed through cross-validation; see Figure 3.6.

How many PLS components should be used?

There are many ways of choosing the number of components for PLS (i.e., the dimensionality of the latent space). Similar to PCA, one can maximize the percent variance explained of original (input X or Y) data. Or, as in regression, one could minimize the mean squared error on predicted (output Y) data. These often (but not always) trade off.

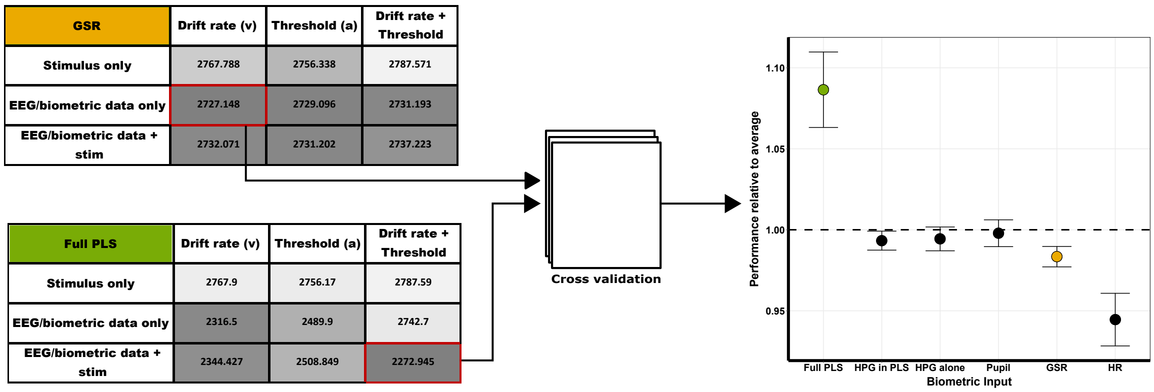

Figure 3.10: Model cross-validated performance relative to competing models. Performance is based on predictive density for each winning model, based on the biometric input type (for computational details see supplementary section 3.5.6). Performance is relative to the average performance for a particular set of test data (i.e., data split 1 though 31) across all competing models (i.e., Pupil versus GSR, versus Full PLS, etc.), to control for the impact of the variability between test data sets. The dotted line at $Y = 1$ indicates average performance, so 1.10 is 10% better performance relative to the average. Error bars are standard deviation for the cross-validated performance for the different biometric input data sets. Each biometric used the winning model as determined above. PLS has the highest performance.

We chose PLS components which best predicted decision parameters in the DDM model. Similar to our model selection above, we performed a model comparison (using DIC score) for each different number of components, in which the components are used to predict both drift v and threshold a scores of the DDM model (see above for example of regression equations). The resulting DIC scores are shown in figure 3.11. As demonstrated, eight components had the lowest DIC score (i.e., best performance), which is what we used for subsequent analyses.

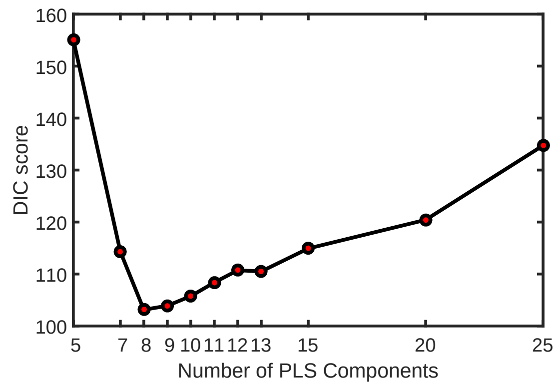

Figure 3.11: Plot of the DIC score for the HDDM model based on different number of components used as regressors (note that for this comparison we use them as regressors in both the drift and threshold parameter). Points indicate each model's DIC score based on the number of PLS components. Note that lower DIC values indicate better model fit. The 8 PLS component model fits the data best.

We are able to extract a decision relevant subspace through PLS dimensionality reduction.

These eight components form a linear covariance space that can both predict decisions and account for the variance in the original biometric features. Note that these components accurately predict measured reaction time (see Figure 3.12). The fact that these components are linear combinations of biometric inputs means that any subsequent analyses on them can be reinterpreted in the original biometric and EEG feature space, retaining interpretability. We caution against over-interpreting the number of components in this case; these eight linear vectors might actually represent and closely approximate only a single nonlinear manifold. Again, while nonlinear directed dimensional reduction techniques exist, we used linear PLS to preserve simplicity and interpretability.

Note that this refers to the use of PLS for full biometrics/EEG. For the HR, pupilometry and GSR (HPG) PLS choice was based easily on 99% variance explained based on very few components.

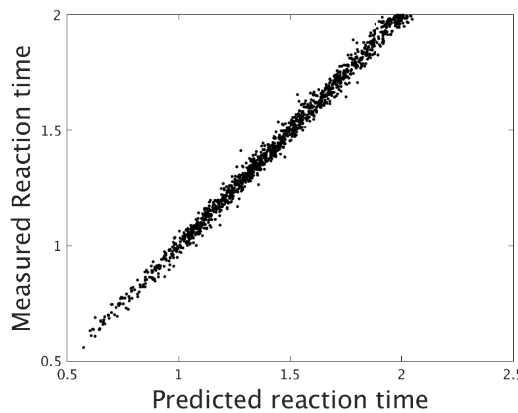

Figure 3.12: Figure demonstrating the performance of the extracted PLS components in the simple linear model from PLS. Extracted components are predictive of decision response time (plotted is measured to predictive response time). Measured reaction time does not exceed 2 seconds, as we employed a cutoff for each trial. $R^2 = 0.9930$.

3.3 Methods

We describe our main methods here. For more details, see supplemental methods in section 3.5.

3.3.1 Participants

We report data from twenty five participants (9 males, 16 females) between the ages of 16 and 38 (mean age = 24) engaged in the study. All participants provided written informed consent and all procedures were approved by the ethics committee of the Institutional Review Board of the University of Minnesota. Participants were compensated with either cash or University course extra credit. All subjects had normal or corrected-to-normal vision and hearing.

3.3.2 Behavioral Paradigm

Participants performed a variation of the random dot motion (RDM) paradigm (Newsome, Britten, & Movshon, 1989). Stimuli were presented using the Psychophysics Toolbox in Matlab (Brainard, 1997), with modified code from G.M. Boynton (University of Washington) (Boynton, n.d.). Participants were seated in a private booth in front of a computer monitor and had access to a keyboard for behavioral responses. Each participant underwent a thresholding procedure to set task difficulty such that individualized accuracy was 75 percent at experiment beginning. The calibration procedure also helped subjects gain familiarity with the task. Once an individualized task difficulty was determined, participants were fit with brain/biometric monitoring equipment and the main experiment began.

On each trial, participants saw a cloud of white dots moving on a black background. Each dot in the cloud appeared only briefly, moved a short distance, then disappeared. Some proportion of dots moved directly left or right (the direction of this subset was the same within each trial). The proportion of uniformly-directed dots is known as the trial's *coherence*. The non-coherent dots moved in random directions. Participants were asked to gauge which direction they perceive the coherent subset of dots to be moving, and to make a corresponding keyboard arrow press as quickly and as accurately as possible.

Each trial began with a five second period of time in which the participant listened to one of three different random sounds (siren, siren+horn, and rain) chosen to induce a range of physiological arousal/alertness levels (see fig(3.3). After the sound stopped, the dot cloud appeared on the screen. The participants had two seconds to make their response. If they responded within the two seconds the dots disappeared and they waited for the next trial to begin. If they did not respond within the two second time interval, the trial was marked as incorrect and the experiment simply continued. If the randomly chosen sound was the siren or the siren+horn together, the participant had a 10 second time period of rain stimulus followed by two seconds of no sound or stimulus intended to allow the participant time to come back to baseline physiology before the next trial began. In the case of the rain sound there was a 2 second inter-trial interval of no sound/no stimulus.

3.3.3 Biometric data collection

We collected a wide variety of brain/body data to facilitate the characterization of brain/body states.

Electroencephalography(EEG)

EEG signals were recorded with a BioSemi ActiveTwo EEG system with ActiView data acquisition software at a sampling rate of 2048 Hz. Participants were fitted with a 64 electrode pin style cap and non-toxic, non-staining electrolytic gel was placed in each pin with an applicator tip in a side to side motion to move hair out of the way facilitating better electrical signal transfer. All electrodes were referenced to two flat electrodes, one placed on each mastoid.

Two flat electrodes were placed, one above and one below the left eye, to facilitate removal of eye movement artifacts. The contact impedance between the EEG electrodes and the scalp were reduced to below +/- 20 mV relative to the CMS/DMR electrodes with the conductive gel.

EEG was preprocessed using the MATLAB-based library Fieldtrip toolbox for EEG and MEG analysis (Oostenveld, Fries, Maris, & Schoffelen, 2011) [1]. EEG data was stored in a Biosemi .bdf files and loaded with Fieldtrip functions. Trial boundaries were defined based on sound stimulus onset, visual stimulus onset, and visual stimulus end. Channels were re-referenced based on the two mastoid electrodes[2]. The function `ft_preprocessing`[3] was used with demeaning to perform baseline correction and detrending to remove linear trends from the data (per trial). A bandpass filter was also applied (between 15-1000 Hz), to remove artifacts[4]. The signal was then downsampled to 512 Hz for further analyses.

Eyeblink removal was accomplished using independent components analysis (ICA) and visual inspection based on comparison with eyeblink channels. Eyeblinks were removed on a per-subject basis[5].

We separated the preprocessed EEG data into pre-visual stimulus and post-stimulus time periods (lined up based on visual stimulus presentation), and concatenated across trials and subjects to create temporal basis functions.

Eye data

Participants were seated approximately 60cm in front of an eye tracking system (Tobii T60XL or Tobii TX300) as they engaged in the task. Eye tracking and pupilometry data were collected and visualized through iMotion's (A/S, n.d.) software interface.

[1] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, the Netherlands. See `http://www.ru.nl/neuroimaging/fieldtrip`

[2] See `https://www.biosemi.com/faq/cms&drl.htm`

[3] See `http://www.fieldtriptoolbox.org/reference/ft_preprocessing`

[4] Following Cavanagh et al. (2011)

[5] Following Frank et al. (2015)

Heart rate (HR)

Each participant was fitted with a Shimmer Optical Pulse Sensing Probe attached to the underside of two fingers via a velco-style wrap. Pulse detection was verified via visual inspection first through both the ShimmerConnect software interface and the iMotions software interface prior to beginning the main experiment.

The heart rate sensor was connected via to a Shimmer3 baseboard situated on a strap on the the participant's wrist. Signal was then transmitted via Bluetooth to a receiver within the booth which connected via USB to the data collection computer.

Galvanic Skin Response (GSR)

Two galvanic skin sensing probes (Electrodermal Resistance Measurement) were attached via velcro-style strips with metallic button snaps to the underside of two fingers of one of the participants hands. Signal was verified visually first through both the ShimmerConnect software interface and the iMotions software interface with a quick startle probe (loud hand clap). The GSR probes were connected to the same Shimmer3 baseboard as the optical pulse sensing probe.

Facial emotion detection

A Logitech HD Pro Webcam C920 was situated above the computer monitor perpendicular to the participants face. The participant booth was well-lit. The web camera feed was connected and visualized within the iMotions interface. The iMotions software then calculated and reported evidence and intensity of facial emotions (i.e., joy, anger, surprise, fear, sadness, contempt, and disgust) as well as individual Action Unit detection on a frame-by-frame analysis of the video feed.

3.3.4 Data cleaning

For the majority of biometrics that were produced by iMotions, we computed the the average value for each pre-stimulus and post-stimulus period, for each trial. For HR and EEG, we performed PLS dimensionality reduction prior to averaging. We removed unmeasured (i.e., NaNs) no-variance values; in such cases, we recorded the trial's average (for that biometric) as NULL. Only trials in which all values were observed were used for producing the PLS weighting components. Trials with missing biometrics were imputed (i.e., using subject averages for that biometric value), and then projected using the previously made PLS weightings. The imputation had a non-significant effect on the weight values (comparison using correlation analysis).

HR data was preprecessed using MATLAB. After extracting photoplethysmography (PPG) data using iMotions software, we performed heart rate estimation on the signal using the open-source MATLAB library ecg-kit (Demski & Llamedo Soria, 2016) and custom written rate-estimation code (see supplemental materials). We then created temporal basis functions for the heart rate signal across trials. This allowed the heart rate on each trial to be represented as a simpler set of basis functions that still capture the majority of the variance, rather than the full time-series.

3.3.5 Partial least squares analysis

Much of dimensionality reduction, such as principal component analysis (PCA), works by mapping a set of measures into a smaller-dimensional space, while preserving most of the variation in the original data. In our case, we want a relationship between our original measures X (an $m \times n$ matrix of measures by trials) and our smaller space of variables Z (a $z \times n$ matrix where $z < m$), where $Z = wX$ (for a set of projection weights w). Importantly, PCA can be treated as an optimization method for finding the set of weights w that maximizes

$Var(wX)$. In order to find this set of weights, PCA uses a linear algebra method called singular value decomposition (SVD) on the correlation matrix of our measures X.

Practically speaking, if we wanted to find the impact of a hidden process on a person's behavior, we could use PCA to find the hidden process Z and use those as regression weights to predict some behavior Y. In other words, we would find natural correlations within the biometric data, find the lower dimensional "subspace" that predicts those correlations, then use those as a measure of the modulatory circuit in a regression.

Since we are interested in the subspace that predicts task-relevant decisions, we are interested in the cross-covariance between the behavior and biometric data. Importantly, if one were to perform PCA independently on X and Y, a subspace for each set of predictors and responses could be found and one could regress against those lower dimensional variables. However, doing PCA on each variable independently potentially throws away the cross-covariance on X and Y. For example, both GSR and HR are impacted by "arousal" - a latent variable. However there are many aspects of measurement that are unrelated to the "arousal" impact on decisions for both HR and GSR.

Similar to PCA, PLS maps recorded measures to lower-dimensional latent space for both X and Y, by treating $Z = wX$ and $U = vY$. However, the weights w and v are found by trying to maximize the prediction of U from Z via linear regression $U = \gamma Z$. In terms of optimization, this means finding the weights that maximize $Cov(wX, vY)^2$, that is, the shared cross-covariance subspace. These weights can then be found by using SVD as above, but on the cross-covariance matrix of X and Y.

PLS performs dimensionality reduction on the cross-covariance, preserving the relationship between X and Y while expressing the idea that X and Y are "contaminated" by errors that are unrelated to the relationship. A related method, canonical correlation analysis (CCA), finds a shared subspace between X and Y observed variables, but instead maximizes

$Corr(wX, vY)^2$. Note that $Var(wX) \times Corr(wX, vY)^2 \times Var(vY) = Cov(wX, vY)^2$, hence PLS is intimately related to CCA. We used PLS because we were concerned with both maximizing the correlation and preserving the original variance in both the X and Y observed spaces.

This subspace found by PLS is interpretable since it is invertable, that is, we can invert any of the steps performed in projecting our data to the low dimensional subspace. For example, in order to interpret the weights γ in the original space we can perform some linear algebra: $\Gamma = w^{-1}\gamma$ where Γ are those regression weights but with respect to X. Because of this, PLS can be used in a variety of ways, including as a method of regularizing (similar to lasso or other prior on weights). It can also be used as an implementation of structural equation modeling (e.g., Kilpatrick and Cahill (2003)). Here, we use it as a way of performing dimensionality reduction on our biometric measures, but preserving their relevant impact on the behaviorally relevant decisions. Since we find the dimensional subspace responsible for behavioral decisions, this constrains our measures to those that are most relevant to the unobserved modulatory circuit's impact on decisions. We can then use the measures projected to this subspace in a non-linear regression needed to model decisions.

We employed PLS using the MATLAB function `plsregress` [6]. We use our notation above where $Z = wX$, $U = vY$ and $U = \gamma Z$ where X is the matrix of trials by features (the predictor variables) and Y is a vector of subject response time (the predicted variable), where w and v are projection matrices (or "loadings"), Z and U are the lower-dimensional results of projecting the original data into the covariance subspace (the scores), and γ are regression weights. In our analysis, Y is a vector of (z-scored) response times (hence we ignore U and v).

We concatentated subject data into a single matrix of trials by features X (where features

[6] see `https://www.mathworks.com/help/stats/plsregress.html`

were all EEG scores, HR scores, and other biometric trial averages for both pre and post task stimuli). We ignored trials with missing data for the creation of the PLS loadings matrix (but imputed values were used to produce scores for use in subsequent analysis). The cleaned X and Y matrices were used as input into the `plsregress` function that outputs both scores and loadings based on the specified number of components.

PLS requires specifying the number of resulting low-dimensional components, which is often selected using either mean-squared error or percent variance explained of original components (which are often traded off against each other). We selected the the number of PLS components by comparing the resulting DIC score when models with different numbers of components were used in the hierarchical drift diffusion regression (see Figure 3.11). Since we planned to use the resulting scores as regressors in the hierarchical drift diffusion model, their performance in this model was the best method for selection. We used DIC scores as a metric for performance. Based on this metric we determined that 8 components had the best performance and was therefore used in the model comparison.

In the initial creation of the loadings w, we used the cleaned feature set with removed trials, and imputed values of biometrics to incorporate otherwise missing trials. Note that imputing values was restricted to only biometric values where a single trial average was missing (e.g., we did not impute missing EEG values). Imputed values were then projected using the original loadings created with non-imputed values, e.g., $Z_{imputed} = wX_{imputed}$ where w is non-imputed. Also note that the MATLAB function rotates most matrices, so in our code, the matrix multiplication is $Z^T_{imputed} = X^T_{imputed}w^T$ (which is identical). Regardless, we then used $Z_{imputed}$ with 8 PLS components as a regressor in all subsequent HDDM model comparisons.

We also performed PLS on the heart rate, galvanic, pupilometry and emotional (facial) data without the EEG, similarly producing a set of lower-dimensional scores (here we chose

4 dimensions based on a similar argument above). This was to allow a comparison on the improved performance with EEG data.

3.3.6 HDDM Model comparisons

In order to further extract brain/body state influence on specific DDM parameters *drift rate* (v) and *threshold* (a) we utilized a Python implementation of the hierarchical drift diffusion model (Wiecki, Sofer, & Frank, 2013). Each model was fit with 10,000 posterior samples with the first 1,000 discarded as burn-in with no thinning. Model strings can be found in supplementary table 3.15. We visually inspected the trace, autocorrelation and the marginal posterior for convergence.

3.3.7 Cross-Validation

In order to assess model fit we generated 31 cross-validated 80:20 splits of the datasets (train:test), and fit the HDDM regression model on each potential Biometric input set (e.g., galvanic vs heart rate vs the 8 PLS components). We made sure that each CV split used the same trials for each Biometric input, e.g., for CV split 1 we used trials (1, 12, 33, ...) for each model (but different values from those trials). This produced MCMC samples for each trial (for each CV split), which were used to produce estimates of the model performances as explained in supplemental info.

We used the `hddm.wfpt.wiener_like` function from the HDDM package (Wiecki, Sofer, & Frank, 2013) to produce estimates from each sample, and averaged them according to equation 3.4. This produces a density value for each of the 31 CV splits. We then averaged across the CV splits for each model. However, since we want to discount the difference in performance due to which trials were selected, we normalized each model's performance based on the other models within the CV split (using vector norm, which divides each by the average performance of all models for that CV split); this produces a performance above average score. This score is plotted in figure 3.10, with 95% confidence intervals (based on the 31 CV split scores). As can be seen, the 8 PLS components produced using the full set of Biometrics performs the best at predicting left-out response time data.

3.4 Discussion

We developed a method to extract decision-relevant biometrics that are reflective of a modulatory process. To accomplish this, we used directed dimensionality reduction (i.e., PLS) and a decision process model (i.e., HDDM). This instantiates how latent decision parameters are impacted by a modulatory state, which is captured by the eight-dimensional space, found through PLS and HDDM regression, which impacts both decision parameters. We confirmed this through both model comparison and cross-validation, showing our prediction performance of the full set of biometric data through PLS is better than the alternatives.

This chapter contributes to our understanding of the impact of biometrics and emotional states on decisions by taking a data-directed rather than a hypothesis-driven view. In past research, biometric measures are generally treated as reflecting a single modulatory process, and detailed experimental work is done to determine which. We instead allow our latent modulatory process to reflect a mixture of observed measures, reflecting the fact that these measures contain both decision-relevant information as well as decision-irrelevant contaminants. This method can therefore be used in a variety of decision tasks which are directly informed by past research.

We use the diffusion model for both its simplicity and success at capturing response times, which is appropriate for our decision task and emotional manipulation. However, in order to capture the impact a modulatory process has on different decision processes multiple types of decisions should be investigated. For example, the gain in a Kalman update (Shadmehr & Mussa-Ivaldi, 2012) or update rate in temporal difference learning (Dayan & Abbott, 2001) can be considered as loci where a modulatory process might impact a decision process. Partial least squares can still be used in these cases, as PLS allows for multidimensional response variables. This does present further complications to make various decision response data be congruent, but is a straightforward extension in many decision tasks (e.g., by using accuracy and response time across multiple tasks). Alternatively, the latent subspace capturing the modulatory process found in one task can be directly compared with another subspace, without directly fitting each subspace across identical response data.

Another limitation, as mentioned in Vandekerckhove (2014), is that the two-step procedure of fitting PLS and then HDDM does not necessarily allow for error propagation in the PLS fitting procedure. There are two main types of errors possible in this case, one being the number of components selected by PLS and another being the error on the fit regression weights, which can limit interpretation of these weights (e.g., due to overfitting). We control for the first by fitting multiple PLS models (with different number of components) and selecting based on HDDM model fit. The second concern can be mitigated by performing a bootstrap of the PLS weights to approximate the error. Provided the error is small (i.e., smaller than the error in the posteriors induced by the HDDM model fit procedure), then this concern can be at least be reasonably mitigated. We performed a post-hoc bootstrap of the PLS weights to compute error, and found the error on the loadings within a reasonable tolerance.

In this chapter, we only present one possible way of using directed dimensionality reduction. Note that while we found eight linear dimensions, this does not mean the dimensional size is meaningfully reflective of any particular computational process. It is entirely possible that the modulatory process spans a nonlinear manifold, and we approximate it here with a linear space. This means our results can be extended through the use of non-linear dimensionality reduction, for example, through deep learning approaches or non-linear embedding (Goodfellow, Bengio, & Courville, 2016). However, the ability to appropriately capture the "correct" latent space must be traded-off with the interpretability of that space. Linearity preserves ease of interpretability, which is appropriate as long as caution is used not to over-interpret.

Emotional states impact decision parameters through modulatory processes, and we demonstrate a method to extract these impacts. This is directly applicable to research that uses biometric data to measure unobserved psychological processes, such as military or pilot training (e.g., Gamble et al. (2018)). Rather than testing individual biometrics, we show how multiple biometric measures can be collected and allow a mixture of them to reflect a latent modulatory space. This modulatory process can then be used to understand the impact of skill training or emotional regulation on individual behavior. It can also inform training, by allowing the investigation of the state of modulatory systems and how they relate to various learning states and performances.

3.5 Supplementary

3.5.1 Biometric preprocessing

For the majority of biometrics that were produced by iMotions, we used trial averages for further analysis (excluding for heart rate and EEG data). Preprocessing included removing any unmeasured (i.e., NaNs) or no variance values, and treating the whole trial's average (for that biometric) as being NULL. Note that we kept with our separation of the pre-stimuli and post-stimuli analysis for the biometrics, producing two trial averages on each trial. For heart rate and EEG, this produces two separate preprocessed vectors of values for each trial;

one before and one after the visual stimuli are initiated.

Only trials in which all values were observed were used for producing the PLS weighting components. Trials with missing biometrics were imputed (i.e., using subject averages for that biometric value), and then projected using the previously made PLS weightings. The imputation had a non-significant effect on the weight values (comparison using correlation analysis).

3.5.2 Heart rate preprocessing

After extracting PPG data using iMotions software, we performed heart rate estimation on the signal, and then created temporal basis functions for the heart rate signal across trials. This allowed the heart rate on each trial to be represented as a simpler set of basis functions, that still capture the majority of the variance, rather than the full time-series.

Heart Rate Estimation

Using the ecg-kit (Demski & Llamedo Soria, 2016)[7], we used the `PPG_pulses_detector` function to estimate peaks of the PPG pulse [8].

`PPG_pulses_detector` estimation is done based on a low-pass differentiator filter. Most parameters were left at default, while refractory period for thresholds was adjusted (default = 150e-3, refract = 300e-3). Parameter choice based on visualization and signal reliability. Since heart rate is known to be within a viable range (e.g., HR cannot go below 10 or above 300 on normal participants), we know the viable range for average heart rate, and used this as a metric to determine appropriate range of parameter values.

[7] From `http://marianux.github.io/ecg-kit/`

[8] From `https://github.com/marianux/ecg-kit/blob/58d4e7e43639da5c0aa6b7cedbb69959f116e755/common/ppg/PPG_pulses_detector.m`

To detect recording dropout, we ran a variance filter over the initial PPG signal; if the variance drops to 0 then no actual measure is being recorded. Filtering was performed using the MATLAB filter function [9] to estimate both mean and variance of signal. Filter uses a rational transfer function, with $\alpha = 0.999$ (where numerator and denominator coefficients b = α and a = $[1, 1 - \alpha]$; see filter.m function for reference). Here we filter an estimate of the mean in order to estimate variance.

Estimating heart rate requires converting peaks to a rate, that is $rate = \frac{events}{time}$. We used an exponential rate filter to estimate instantaneous heart rate.

Exponential rate filter

Following Karonen (2014), we used an exponentially weighted rolling filter to estimate the instantaneous rate of heart beats. Rates are defined as the number of events over a time period, that is $\lambda = \frac{N}{\Delta t}$ where N is total number of events and Δt is some time period. If the rate is constant than you can simply estimate. However to produce an estimate of a changing rate a filtering approach is preferred.

To produce an exponentially weighted estimate, we update our previous rate estimate by multiplying it by an exponentially decreasing weight based on when the last event occurred (where events are the estimated peaks). This means if an event has not happened recently then we downweight our estimate. Estimates of the rates can only be updated at events.

In mathematical notation, assume we want to estimate the true rate λ with an estimate λ^*, where we have events occuring at times $t_1, t_2, ..., t_n$ (where t_0 is the initial time). Again, if the rate is constant we estimate simply by:

$$\lambda^* = \frac{N}{\Delta t}$$

[9] See https://www.mathworks.com/help/matlab/ref/filter.html

where $\Delta t = t - t_0$ for the current time t. This can be stated more generally using the Dirac delta function δ:

$$\lambda^* = \frac{\int_{t_0}^{\tau} \sum_{i=1}^{N} \delta(\tau - t_i) d\tau}{\int_{t_0}^{\tau} 1 d\tau}$$

While this is equivalent to our estimate above, it demonstrates how we can estimate rate via an instantaneous measure of events and time. $\sum_{i=1}^{N} \delta(\tau - t_i)$ teasures the instantaneous rate as which events increase at τ, which the integral then counts up over Δt, while the denominator measures how time increases (which is constant).

Since our rate changes, we want to limit the window over which we estimate rate by using a weighting function $w(\tau)$ that downweights more distant times.

$$\lambda^* = \frac{\int_{t_0}^{\tau} w(\tau) \sum_{i=1}^{N} \delta(\tau - t_i) d\tau}{\int_{t_0}^{\tau} w(\tau) d\tau}$$

Choosing an exponentially decaying function $w(\tau) = e^{k(\tau - t)}$ for some decay rate $0 < k < 1$, allows us to simplify this equation.

$$\begin{aligned} \lambda^* &= \frac{\int_{t_0}^{\tau} e^{k(\tau-t)} \sum_{i=1}^{N} \delta(\tau - t_i) d\tau}{\int_{t_0}^{\tau} e^{k(\tau-t)} d\tau} \\ &= \frac{\int_{t_0}^{\tau} \sum_{i=1}^{N} e^{k(\tau-t)} \delta(\tau - t_i) d\tau}{-\frac{1}{k}(e^{k(t_0-t)} - 1)} \\ &= \frac{\sum_{i=1}^{N} e^{k(t_i-t)}}{-\frac{1}{k}(e^{k(t_0-t)} - 1)} \\ &= \frac{k \sum_{i=1}^{N} e^{k(t_i-t)}}{1 - e^{k(t_0-t)}} \end{aligned}$$

Note that for any function $f(a)$, $\int_{a-\epsilon}^{a+\epsilon} f(a)\delta(x - a)dx = f(a)$ for $\epsilon > 0$, which allows us to remove the integral in the numerator. In addition, if we treat t_0 as essentially $-\infty$ (i.e., that the total observation time is much larger than the decay weight), then we can remove the denominator and treat our estimate as:

$$\lambda^* = k \sum_{i=1}^{N} e^{k(t_i - t)}$$

This estimate can now be turned into a recursive update equation via a similar method to exponentially weighted averages, by assuming we have some earlier estimate $\lambda^*(t')$ where t' is the last observed event, we can compute the current weighted average event rate:

$$\lambda^*(t) = e^{k(t'-t)}\lambda^*(t') \tag{3.3}$$

If an event occurs at time t, then instead we have:

$$\lambda^*(t) = k + e^{k(t'-t)}\lambda^*(t')$$

Our update rule for an exponentially weighted rate, then, is, when we observe an event i at time t_i where t_{i-1} is the previous event:

$$\lambda_i^* \leftarrow k + e^{k(t_{i-1}-t_i)}\lambda_{i-1}^*$$

We can then use equation 3.3 to produce an instantaneous estimate of the rate.

We can also perform two corrections to this estimate. One is to correct for an initial bias of our initial start time, and add back in the denominator which we initially removed to equation 3.3:

$$\lambda_{corrected}^* = \frac{e^{k(t'-t)}\lambda^*(t')}{1 - e^{k(t_0 - t)}}$$

where t_0 is the first measured time.

Second, we can also smooth our estimate by performing an exponentially weighted estimate of λ^* (here λ^{**}).

$$\lambda^{**}(t) = \frac{\int_{-\infty}^{t} \lambda^*(\tau) e^{k_2(\tau - t)} d\tau}{\int_{-\infty}^{t} e^{k_2(\tau - t)} d\tau}$$

which can be calculated in a stepwise fashion:

$$\lambda_i^{**} \leftarrow W(t_{i-1} - t_i)\lambda_{i-1}^{*} + e^{-k_2(t_{i-1}-t_i)}\lambda_{i-1}^{**}$$
$$\lambda_i^{*} \leftarrow k_1 + e^{-k_1(t_{i-1}-t_i)}\lambda_{i-1}^{*}$$

Again where k_1 and k_2 are decay rates between 0 and 1, and:

$$W(t_{i-1} - t_i) = \begin{cases} k_2 \frac{e^{-k_2(t_{i-1}-t_i)} - e^{-k_1(t_{i-1}-t_i)}}{k_1 - k_2}, & \text{if } k_1 \neq k_2 \\ k(t_{i-1} - t_i)e^{-k(t_{i-1}t_i)} & \text{if } k_1 = k_2 = k \end{cases}$$

Then we can similarly do a correction over this estimate as in $\lambda^{*}_{corrected}$ above.

$$\lambda^{**}_{corrected}(t) = \frac{\lambda^{**}(t)}{1 - S(t - t_0)}$$

where:

$$S(\Delta t) = \begin{cases} \frac{k_1 e^{-k_2 \Delta t} - k_2 e^{-k_1 \Delta t}}{k_1 - k_2} & \text{if } k_1 \neq k_2 \\ (1 + k\Delta t)e^{-k\Delta t} & \text{if } k_1 = k_2 = k \end{cases}$$

Derivation for the exponential rate filter originally from Karonen (2014).

Smoothing estimates for missed events

Since signal dropout occured (due to e.g., the heart rate monitor being disrupted by subject movements), we must incorporate missed events into our estimate. Points without signal were considered "no events" points during filtering, and for the above filtering we did not include them (simply treating the time points as no event having occurred). However this is a poor estimate. At any time period where there was no event (provided it was significantly smaller than a trial), we used a smoothing spline to provide us with a continuous estimate, based on the estimated rate at those events that had occurred.

First we downsampled the filtered rate estimates from λ^{**}; since events occur at a certain time frequency, we cannot get a good estimate of the rate in times between events. So we downsampled our estimated rate to be on the order of events.

In order to deal with missing signal, we used a smoothing spline via the csaps function in MATLAB. The csaps function uses a cubic spline estimate, using the event times as knot points.

This produces a smoothed signal. However instead of using the full estimate time-series for further analyses, we employed a temporal basis method that explains 99% of the variance.

Heart Rate temporal basis functions

We produced temporal basis functions for the heart rate time series. We took our time series and stacked all trials for all subjects into an M X T matrix X (where M = number of subjects times number of trials, and T = total number of samples per trial). Note that we will have two matrices for pre-stimuli and post-stimuli heart rate.

We can perform principal components analysis on X (here we use singular value decomposition on the covariance C of X) to find a basis set of vectors that are length T that find the lower-dimension representation of the time series based on the trial by trial covariances.

We projected down to produce a lower dimension set of values for each trial that captured 99% of the variance of each trial's time-series: 3 dimensions for the pre-stimuli time and 2 dimensions for the post-stimuli.

These scores were then used in subsequent analysis similar to trial averages to other biometric values.

3.5.3 EEG preprocessing

EEG was preprocessed using the MATLAB based library Fieldtrip toolbox for EEG/MEG-analysis (Oostenveld, Fries, Maris, & Schoffelen, 2011)[10]. EEG data was stored in a Biosemi .bdf files and loaded with fieldtrip functions. Trials were defined based on sound stimulus onset, visual stimulus onset, and visual stimulus end. Channels were rereferenced based on the two mastoid electrodes [11]. The function `ft_preprocessing` [12] was used with demeaning to perform baseline correction and detrending to remove linear trends from the data (per trial). A bandpass filter was also applied (between 15-1000 Hz), to remove artifacts (following (Cavanagh et al., 2011)). The signal was then downsampled to 512 Hz (primarily for computational resource reasons).

Eyeblink removal was accomplished using independent components analysis (ICA) and visual inspection based on comparison with eyeblink channels. Eyeblinks were removed on a per-subject basis (following Frank et al. (2015)).

Afterward, we separated the preprocessed EEG data into pre-visual stimulus and post-stimulus time periods (lined up based on visual stimulus presentation), and concatenated across trials and subjects to create temporal basis functions.

EEG temporal basis functions

Temporal basis functions were created using principal components analysis (PCA) to reduce the dimensionality for the time-series for each EEG channel. Separate temporal bases were created for the pre- and post-stimulus time periods.

[10] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, the Netherlands. See `http://www.ru.nl/neuroimaging/fieldtrip`

[11] See `https://www.biosemi.com/faq/cms&drl.htm`

[12] See `http://www.fieldtriptoolbox.org/reference/ft_preprocessing`

Temporal basis functions were created by performing SVD on the channel time-series covariance matrix, with the number of components explaining over 90% of the variance (600 components were used for both pre and post stimulus times). Each component represents a weighted average across the time-series for a single channel.

The EEG data was then rearranged so that each trial data could be concatenated with biometric, producing a trial X variable matrix (trials across all subjects) for the partial least squares analysis.

3.5.4 Partial least squares analysis

As explained above, we employed partial least squares as a method of directed dimensionality reduction using the MATLAB function `plsregress`[13]. Note MATLAB uses the SIMPLS algorithm Jong (1993). We use our notation above where $Z = wX$, $U = vY$ and $U = \gamma Z$ where X is the matrix of trials by features (the predictor variables) and Y is a vector of subject response time (the predicted variable), where w and v are projection matrices (or loadings), Z and U are the lower-dimensional results of projecting the original data into the covariance subspace (the scores), and γ are regression weights. For us Y is a vector of (z-scored) response times (hence we ignore U and v).

Subject data was concatenated into a single matrix of trials by features X (where features were all EEG scores, HR scores, and other biometric trial averages for both pre and post task stimuli, see Figure 3.13). Trials with missing data were ignored for the creation of the PLS loadings matrix (but imputed values were used to produce scores for use in subsequent analysis). The cleaned X and Y matrices were used as input into the `plsregress` function that outputs both scores and loadings based on the specified number of components.

[13]See `https://www.mathworks.com/help/stats/plsregress.html`

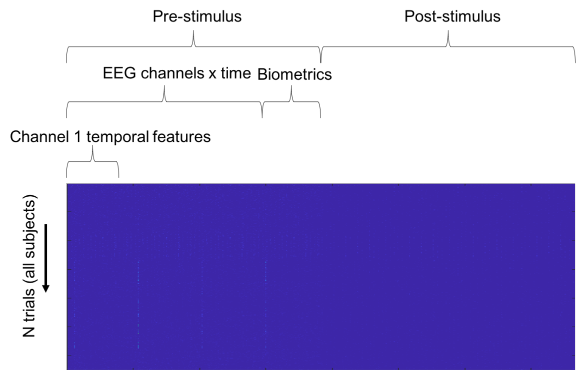

Figure 3.13: Figure demonstrating the way subject data was concatenated. This produces the X matrix below. Note that this figure actually shows the transpose of the X matrix below.

To specify the number of components for PLS, we chose based on their DIC score when used in the hierarchical drift diffusion regression (see figure 3.11). PLS requires specifying the number of resulting low-dimensional components, which is often selected using either mean-squared error or percent variance explained of the original components (which are often traded off against each other). Since we planned to use the resulting scores as regressors in the hierarchical drift diffusion model, their performance in this model was the best method for selection. We used DIC scores as a metric for performance. Based on this metric we determined that 8 components had the best performance and was therefore used in the model comparison.

In the initial creation of the loadings w we used the cleaned feature set with removed trials, however we used imputed values of biometrics to incorporate otherwise missing trials. Note that imputing values was restricted to only biometric values where a single trial average was missing (e.g., we did not impute missing EEG values). Imputed values were then projected using the original loadings created with non-imputed values, e.g., $Z_{imputed} = wX_{imputed}$ where w is non-imputed. Also note that the MATLAB function rotates most matrices, so in our code the matrix multiplication is $Z^T_{imputed} = X^T_{imputed}w^T$

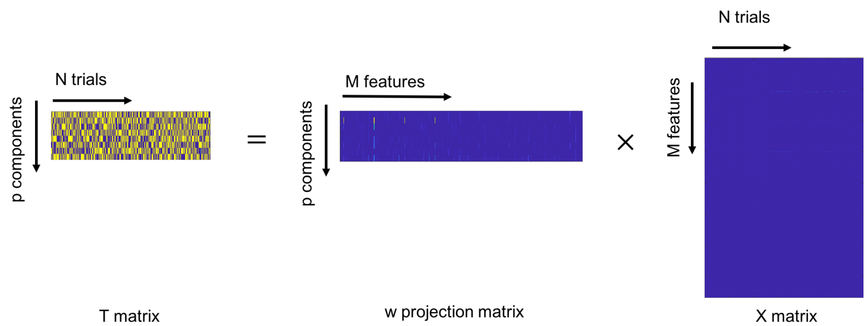

Figure 3.14: Figure demonstrating relationship between original features matrix X and the resulting scores. Here $Z = wX$, where w is a projection matrix that reduces the feature space down to the p components (in our case, 8).

(which is identical). Regardless, we then used $Z_{imputed}$ with 8 PLS components as a regressor in all subsequent HDDM model comparisons.

We also performed PLS on the heart rate, galvonic, pupilometry and emotional (facial) data without the EEG, similarly producing a set of lower-dimensional scores (here we chose 4 dimensions based on a similar argument above). This was to allow a comparison on the improved performance with EEG data.

3.5.5 Hierarchical Bayesian estimation of the Drift Diffusion Model

Full model strings are specified in table 3.15, table 3.16, and table 3.17.

Input Data	Model N	N Model Inputs	Model String	DIC
GSR	1	3	(data, depends_on={'a': 'stim'})	2767.788
GSR	2	3	(data, depends_on={'v': 'stim'})	2756.338
GSR	3	6	(data, depends_on={'a': 'stim', 'v': 'stim'})	2787.571
GSR	4	2	(data, 'a ~ pre_gsr + post_gsr')	2727.148
GSR	5	2	(data, 'v ~ pre_gsr + post_gsr')	2729.096
GSR	6	4	(data, {'a ~ pre_gsr + post_gsr', 'v ~ pre_gsr + post_gsr'})	2731.193
GSR	7	6	(data, 'a ~ pre_gsr:C(stim) + post_gsr:C(stim)')	2732.071
GSR	8	6	(data, 'v ~ pre_gsr:C(stim) + post_gsr:C(stim)')	2731.202
GSR	9	12	(data, {'a ~ pre_gsr:C(stim) + post_gsr:C(stim)', 'v ~ pre_gsr:C(stim) + post_gsr:C(stim)'})	2737.223
Pupil	1	3	(data, depends_on={'a': 'stim'})	2767.192
Pupil	2	3	(data, depends_on={'v': 'stim'})	2756.614
Pupil	3	6	(data, depends_on={'a': 'stim', 'v': 'stim'})	2787.783
Pupil	4	4	(data, 'a ~ pre_pL + pre_pR + post_pL + post_pR')	2690.348
Pupil	5	4	(data, 'v ~ pre_pL + pre_pR + post_pL + post_pR')	2686.52
Pupil	6	8	(data, {'a ~ pre_pL + pre_pR + post_pL + post_pR', 'v ~ pre_pL + pre_pR + post_pL + post_pR'})	2670.922
Pupil	7	12	(data, 'a ~ pre_pL:C(stim) + pre_pR:C(stim) + post_pL:C(stim) + post_pR:C(stim)')	2688.603
Pupil	8	12	(data, 'v ~ pre_pL:C(stim) + pre_pR:C(stim) + post_pL:C(stim) + post_pR:C(stim)')	2692.851
Pupil	9	24	(data, {'a ~ pre_pL:C(stim) + pre_pR:C(stim) + post_pL:C(stim) + post_pR:C(stim)', 'v ~ pre_pL:C(stim) + pre_pR:C(stim) + post_pL:C(stim) + post_pR:C(stim)'})	2679.119
HR	1	3	(data, depends_on={'a': 'stim'})	2767.317
HR	2	3	(data, depends_on={'v': 'stim'})	2757.252
HR	3	6	(data, depends_on={'a': 'stim', 'v': 'stim'})	2788.637
HR	4	4	(data, 'a ~ HR_imputedscores1 + HR_imputedscores2 + HR_imputedscores3 + HR_imputedscores4')	2731.711
HR	5	4	(data, 'v ~ HR_imputedscores1 + HR_imputedscores2 + HR_imputedscores3 + HR_imputedscores4')	2730.797
HR	6	8	(data, {'a ~ HR_imputedscores1 + HR_imputedscores2 + HR_imputedscores3 + HR_imputedscores4', 'v ~ HR_imputedscores1 + HR_imputedscores2 + HR_imputedscores3 + HR_imputedscores4'})	2735.981
HR	7	12	(data, 'a ~ HR_imputedscores1:C(stim) + HR_imputedscores2:C(stim) + HR_imputedscores3:C(stim) + HR_imputedscores4:C(stim)')	2738.902
HR	8	12	(data, 'v ~ HR_imputedscores1:C(stim) + HR_imputedscores2:C(stim) + HR_imputedscores3:C(stim) + HR_imputedscores4:C(stim)')	2729.47

Figure 3.15: Full model strings for all HDDM models. Number of model inputs highlighted. DIC values reported from each model ran

Input Data	Model N	N Model Inputs	Model String	DIC
HR	9	24	(data, {'a ~ HR_imputedscores1:C(stim) + HR_imputedscores2:C(stim) + HR_imputedscores3:C(stim) + HR_imputedscores4:C(stim)', 'v ~ HR_imputedscores1:C(stim) + HR_imputedscores2:C(stim) + HR_imputedscores3:C(stim) + HR_imputedscores4:C(stim)'})	2737.925
HPG alone	1	3	(data, depends_on={'a': 'stim'})	2768.002
HPG alone	2	3	(data, depends_on={'v': 'stim'})	2757.356
HPG alone	3	6	(data, depends_on={'a': 'stim', 'v': 'stim'})	2787.46
HPG alone	4	11	(data, 'a ~ pre_pL + pre_pR + pre_gsr + pre_hr1 + pre_hr2 + pre_hr3 + post_pL + post_pR + post_gsr + post_hr1 + post_hr2')	2696.15
HPG alone	5	11	(data, 'v ~ pre_pL + pre_pR + pre_gsr + pre_hr1 + pre_hr2 + pre_hr3 + post_pL + post_pR + post_gsr + post_hr1 + post_hr2')	2694.15
HPG alone	6	22	(data, {'a ~ pre_pL + pre_pR + pre_gsr + pre_hr1 + pre_hr2 + pre_hr3 + post_pL + post_pR + post_gsr + post_hr1 + post_hr2', 'v ~ pre_pL + pre_pR + pre_gsr + pre_hr1 + pre_hr2 + pre_hr3 + post_pL + post_pR + post_gsr + post_hr1 + post_hr2'})	2682.872
HPG alone	7	33	(data, 'a ~ pre_pL:C(stim) + pre_pR:C(stim) + pre_gsr:C(stim) + pre_hr1:C(stim) + pre_hr2:C(stim) + pre_hr3:C(stim) + post_pL:C(stim) + post_pR:C(stim) + post_gsr:C(stim) + post_hr1:C(stim) + post_hr2:C(stim)')	2708.366
HPG alone	8	33	(data, 'v ~ pre_pL:C(stim) + pre_pR:C(stim) + pre_gsr:C(stim) + pre_hr1:C(stim) + pre_hr2:C(stim) + pre_hr3:C(stim) + post_pL:C(stim) + post_pR:C(stim) + post_gsr:C(stim) + post_hr1:C(stim) + post_hr2:C(stim)')	2713.951
HPG alone	9	66	(data, {'a ~ pre_pL:C(stim) + pre_pR:C(stim) + pre_gsr:C(stim) + pre_hr1:C(stim) + pre_hr2:C(stim) + pre_hr3:C(stim) + post_pL:C(stim) + post_pR:C(stim) + post_gsr:C(stim) + post_hr1:C(stim) + post_hr2:C(stim)', 'v ~ pre_pL:C(stim) + pre_pR:C(stim) + pre_gsr:C(stim) + pre_hr1:C(stim) + pre_hr2:C(stim) + pre_hr3:C(stim) + post_pL:C(stim) + post_pR:C(stim) + post_gsr:C(stim) + post_hr1:C(stim) + post_hr2:C(stim)'})	2716.127
HPG in PLS	1	3	(data, depends_on={'a': 'stim'})	2768.058
HPG in PLS	2	3	(data, depends_on={'v': 'stim'})	2757.569
HPG in PLS	3	6	(data, depends_on={'a': 'stim', 'v': 'stim'})	2787.974
HPG in PLS	4	4	(data, 'a ~ HPG_imputedscores1 + HPG_imputedscores2 + HPG_imputedscores3 + HPG_imputedscores4')	2704.166
HPG in PLS	5	4	(data, 'v ~ HPG_imputedscores1 + HPG_imputedscores2 + HPG_imputedscores3 + HPG_imputedscores4')	2704.995
HPG in PLS	6	8	(data, {'a ~ HPG_imputedscores1 + HPG_imputedscores2 + HPG_imputedscores3 + HPG_imputedscores4', 'v ~ HPG_imputedscores1 + HPG_imputedscores2 + HPG_imputedscores3 + HPG_imputedscores4'})	2693.772

Figure 3.16: Full model strings for all HDDM models. Number of model inputs highlighted. DIC values reported from each model ran

Input Data	Model N	N Model Inputs	Model String	DIC
HPG in PLS	7	12	(data, 'a ~ HPG_imputedscores1:C(stim) + HPG_imputedscores2:C(stim) + HPG_imputedscores3:C(stim) + HPG_imputedscores4:C(stim)')	2712.004
HPG in PLS	8	12	(data, 'v ~ HPG_imputedscores1:C(stim) + HPG_imputedscores2:C(stim) + HPG_imputedscores3:C(stim) + HPG_imputedscores4:C(stim)')	2707.51
HPG in PLS	9	24	(data, {'a ~ HPG_imputedscores1:C(stim) + HPG_imputedscores2:C(stim) + HPG_imputedscores3:C(stim) + HPG_imputedscores4:C(stim)', 'v ~ HPG_imputedscores1:C(stim) + HPG_imputedscores2:C(stim) + HPG_imputedscores3:C(stim) + HPG_imputedscores4:C(stim)'})	2701.889
Full PLS	1	3	(data, depends_on={'a': 'stim'})	2767.9
Full PLS	2	3	(data, depends_on={'v': 'stim'})	2756.17
Full PLS	3	6	(data, depends_on={'a': 'stim', 'v': 'stim'})	2787.59
Full PLS	4	8	(data, 'a ~ PLS_Scores1 + PLS_Scores2 + PLS_Scores3 + PLS_Scores4 + PLS_Scores5 + PLS_Scores6 + PLS_Scores7 + PLS_Scores8')	2316.5
Full PLS	5	8	(data, 'v ~ PLS_Scores1 + PLS_Scores2 + PLS_Scores3 + PLS_Scores4 + PLS_Scores5 + PLS_Scores6 + PLS_Scores7 + PLS_Scores8')	2489.9
Full PLS	6	16	(data, {'a ~ PLS_Scores1 + PLS_Scores2 + PLS_Scores3 + PLS_Scores4 + PLS_Scores5 + PLS_Scores6 + PLS_Scores7 + PLS_Scores8', 'v ~ PLS_Scores1 + PLS_Scores2 + PLS_Scores3 + PLS_Scores4 + PLS_Scores5 + PLS_Scores6 + PLS_Scores7 + PLS_Scores8'})	2742.7
Full PLS	7	24	(data, 'a ~ PLS_Scores1:C(stim) + PLS_Scores2:C(stim) + PLS_Scores3:C(stim) + PLS_Scores4:C(stim) + PLS_Scores5:C(stim) + PLS_Scores6:C(stim) + PLS_Scores7:C(stim) + PLS_Scores8:C(stim)')	2344.427
Full PLS	8	24	(data, 'v ~ PLS_Scores1:C(stim) + PLS_Scores2:C(stim) + PLS_Scores3:C(stim) + PLS_Scores4:C(stim) + PLS_Scores5:C(stim) + PLS_Scores6:C(stim) + PLS_Scores7:C(stim) + PLS_Scores8:C(stim)')	2508.849
Full PLS	9	48	(data, {'a ~ PLS_Scores1:C(stim) + PLS_Scores2:C(stim) + PLS_Scores3:C(stim) + PLS_Scores4:C(stim) + PLS_Scores5:C(stim) + PLS_Scores6:C(stim) + PLS_Scores7:C(stim) + PLS_Scores8:C(stim)', 'v ~ PLS_Scores1:C(stim) + PLS_Scores2:C(stim) + PLS_Scores3:C(stim) + PLS_Scores4:C(stim) + PLS_Scores5:C(stim) + PLS_Scores6:C(stim) +	2272.945

Figure 3.17: Full model strings for all HDDM models. Number of model inputs highlighted. DIC values reported from each model ran

3.5.6 Computing cross-validation evidence

Our goal is to compute the cross validated evidence for each of our models. In particular:

$$P(Data_{test}|Data_{train})$$

for each model. $Data_{test}$ draws the reaction time for each trial y_i, so:

$$\begin{aligned} P(Data_{test}|Data_{train}) =& \prod_{i=1}^{M} P(y_i|Data_{train}) \\ =& exp(\frac{1}{M}\sum_{i=1}^{M} log\ P(y_i|Data_{train}))^M \end{aligned}$$

Where $P(y_i|Data_{train})$ is estimated from the samples. We can drop the M exponent if we want an average estimate per trial rather than overall evidence.

We use the function `hddm.wfpt.wiener_like` from the HDDM package (Wiecki, Sofer, & Frank, 2013), which when given the DDM parameters of a, v, and t (here we call all parameters θ), we can output the log density for the reaction time and accuracy (called y_i for the ith trial). We call this log density function for the wiener distribution: $log\, p(y_i|\theta_{ji}) = z_{ij}$, where $\theta_{ji} = \bar{\beta}_j \bar{x}_i + \alpha_{ij}$. Since we are using a sampling algorithm, we have j samples and i data points (for a given cross-validation). Remember that each i data point is for a given trial for a given subject. Note that the specific equation for a given θ depends on the model being fit; the beta weights are from the j samples, while the X refers to any biometric data on the ith trial. The alpha refers to the regression intercept, which is fit for subject (so depends on the data trial i).

We want to compute the evidence, so marginalize across the parameters:

$$\begin{aligned} P(y_{new}|Data_{train}) =& \sum_{\theta} P(y|\theta)P(\theta|Data_{train}) \\ \approx& \frac{1}{N}\sum_{i=1}^{N} P(y_i|\theta_j) \end{aligned}$$

where $\theta_i \approx P(\theta|Data_{train})$. This approximation is the estimate driven from the samples. This means to compute our estimate of the probability of data for a trial, we have:

$$P(y_i|Data_{train}) = \frac{1}{N}\sum_{i=1}^{N} exp(z_{i,j})$$

So, we obtain:

$$\begin{aligned} P(y_{test}|Data_{train}) &= exp(\frac{1}{M}\sum_{i=1}^{M} log\ P(y_i|Data_{train})) \\ &= exp(\frac{1}{M}\sum_{i=1}^{M} log\ (\frac{1}{N}\sum_{i=1}^{N} exp(z_{i,j}))) \end{aligned}$$

For each cross-validation, for each trial. To compute the expected evidence across cross-validation s (let each cross-validation be indexed by k), we have:

$$\begin{aligned} \bar{P}(y^k_{test}|Data^k_{train}) &= E_k[exp(\frac{1}{M}\sum_{i=1}^{M} log\ (\frac{1}{N}\sum_{i=1}^{N} exp(z_{i,j})))] \\ &= \frac{1}{k}\sum_{k}[exp(\frac{1}{M}\sum_{i=1}^{M} log\ (\frac{1}{N}\sum_{i=1}^{N} exp(z_{i,j})))] \end{aligned} \tag{3.4}$$

Functionally we compute z_{ij} for each trial i, sample j, cross-validation k, and model, and then collapse appropriately across each index to compare models on their evidence.

We generated 31 cross-validated 80:20 splits of the datasets (train:test), and fit the HDDM regression model on each potential Biometric input set (e.g., galvanic vs heart rate vs the 8 PLS components). We made sure that each CV split used the same trials for each Biometric input, e.g., for CV split 1 used trials (1, 12, 33, ...) for each model (but different values from those trials). This produced MCMC samples for each trial (for each CV split), which were used to produce estimates of the model performances as explained above.

Again, we used the `hddm.wfpt.wiener_like` function to produce estimates from each sample, and averaged them according to equation 3.4. This produces a density value for each of the 31 CV splits. We then averaged across the CV splits for each model. However, since we want to discount the difference in performance due to which trials were selected, we normalized each model's performance based on the other models within the CV split (using vector norm, that is divide each by the average performance of all models for that CV split); this produces a performance above average score. This score is plotted in figure 3.10, with 95% confidence intervals (based on the 31 CV split scores). As can be seen, the 8 PLS components produced using the full set of Biometrics performs the best at predicting left-out response time data.

Interlude

In order to allocate effort and time appropriately, meta-cognitive processes must forecast the future impact of such allocation. In Chapter 4, we investigate the role of subjective confidence as an example of such a forecast, where confidence reflects the reliability of information, and allows adaptive setting of information integration, adjusting the time allocated to a task in response to the information reliability.

Chapter 4

Confidence reflects internal information gain

4.1 Introduction

Flexible decision making requires the ability to integrate information across time while dealing with uncertainty. This uncertainty comes both from external sources (Green & Swets, 1966) and internal sources such as attentional fluctuations (Denison, Adler, Carrasco, & Ma, 2018). In perceptual decision making much of this uncertainty is reflected in subjective confidence (Mamassian, 2016). While subjective confidence is a salient part of decisions, there are various ways in which reported confidence might interact with decision processes. What is the role of confidence on information integration in decision making?

Many models of decision making attempt to account for the choice and timing of decision through the accumulation of information up to a decision threshold. We investigate the meta-cognitive role confidence has in monitoring decision processes, using systems identification methods. This requires generalizing standard integration models to include time-varying threshold and weights on information. We believe confidence has a role in reflecting and monitoring information gain, which can be used by the system to adaptively set decision criteria.

4.1.1 Background

Confidence

Confidence refers to a metacognitive, subjective sense a person has concerning their own abilities or performance with respect to some task (Yeung & Summerfield, 2012; Mamassian, 2016). While the experience of confidence is salient, the role confidence might play in decision processes is unclear. Much of this is due to the relationship between confidence and performance in tasks. For perceptual tasks, it is well-established that confidence increases with stronger stimuli (Lange, Gaal, Lamme, & Dehaene, 2011; Peirce & Jastrow, 1884) and with the time given to sample stimuli (Vickers, Smith, Burt, & Brown, 1985). Biases in confidence, such that participants are often overconfident in correct trials. This bias appears related to physiological arousal state (Jönsson, Olsson, & Olsson, 2005), are also commonly found. This "overconfidence" phenomenon relates to the much debated "hard-easy" effect (Juslin, Winman, & Olsson, 2000; Merkle, 2009; Moore & Healy, 2008), in that participants are more overconfident in harder problems and exhibit under-confidence on easier problems (Drugowitsch, Moreno-Bote, & Pouget, 2014).

Confidence is measured in a variety of ways for both human and nonhuman participants[1]. Many methods assess the implications of confidence, such as post-decision wagering in which a participants places a bet on their choice (Persaud, McLeod, & Cowey, 2007). Other "objective" measures of confidence include amount of time waiting for a reward (Kepecs, Uchida, Zariwala, & Mainen, 2008), or having a "safe bet" option that is always rewarded (Kiani & Shadlen, 2009). In these cases, confidence is presumed to represent a participant's uncertainty concerning a rewarding outcome, in that they will act riskier if they are more confident in their decision. While these measures have the benefit of re-

[1] For a review see Kepecs and Mainen (2014)

quiring only overt behavior, and so can be used across non-human animals, it is not clear that they directly correspond to subjective confidence (Adler & Ma, 2018). By contrast, the simplest measure of confidence is subjective report, that is, asking a person how confident they are in some decision, generally on an integer scale (Denison, Adler, Carrasco, & Ma, 2018; Kiani, Corthell, & Shadlen, 2014; Koriat, 2012; Peirce & Jastrow, 1884; Kunimoto, Miller, & Pashler, 2001; Sanders, Hangya, & Kepecs, 2016; Jönsson, Olsson, & Olsson, 2005; Fleming, Huijgen, & Dolan, 2012; Fleming, Weil, Nagy, Dolan, & Rees, 2010; Lange, Gaal, Lamme, & Dehaene, 2011; Charles, Van Opstal, Marti, & Dehaene, 2013; Bang et al., 2014; Adler, 2018, January). Subjective reports are beneficial in that they directly assesses the introspective "sense of knowing" (Brown, 1991) that we are most interested in, and make far fewer assumptions concerning the role confidence might play in decision making.

Many models instantiate confidence as a variable measuring uncertainty concerning decisions or perceptions, and so are generally specified in standard probabilistic terms (Mamassian, 2016). A common statistical view, for example, relates subjective confidence to statistical confidence (Sanders, Hangya, & Kepecs, 2016), stating that subjective confidence refers to the posterior probability of being correct (Adler, 2018, January; Drugowitsch, Moreno-Bote, & Pouget, 2014; Meyniel, Sigman, & Mainen, 2015; Pouget, Drugowitsch, & Kepecs, 2016). In these cases, confidence is a simple reflection of choice-relevant decision variables, such that they may be read directly from the decision process itself. This is sensible for post-decision confidence, which might allow for additional distortions to explain the previously mentioned biases in confidence (Sanders, Hangya, & Kepecs, 2016), but is less so for confidence that might occur before or during the decision. Similarly, while relating confidence to objective sensory information permits standard psychometric methods, it ignores critical internal sources of uncertainty, such as those driven by attention

(Denison, Adler, Carrasco, & Ma, 2018). This suggests that confidence is a metacognitive variable that tracks multiple sources of evidence reliability, both internal and external (Yeung & Summerfield, 2012). In these cases, confidence cannot be simply read from decision variables, but instead is allowed an active role in shaping choice decisions. Formal models of information integration can allow us to specify how this tracking occurs, and what result it might have on the decision process.

Information integration models

Information integration in humans and animals is often modeled as a sequential sampling process. In a sequential sampling model (SSM), decisions are made by accumulating information that reflects one of multiple choices until a threshold is reached indicating a response. These models allow the inference of response time based on their sequential evidence accumulation, as in many perceptual tasks. A common SSM is the drift diffusion model (DDM), a highly reliable and popular model of human perceptual decisions made between two choices (see Ratcliff, Smith, Brown, and McKoon (2016) for a recent review). Consider a decision between two choices based on noisy incoming information. At any point, you must decide both to stop collecting information and which choice to make. Choice-relevant information is assumed to be represented as a decision variable v_t at each time point, with decisions made once the cumulative information $I = \sum v_t$ passes a boundary, $|I| < a$ (with the sign of I often representing each choice). This model largely fits with human reaction time data (Ratcliff & McKoon, 2008; Voss, Rothermund, & Voss, 2004) and neural evidence of information accumulation (Bogacz & Bogacz, 2007), and it can be derived from normative stopping rules (Ratcliff & Smith, 2004).

A common extension of the diffusion model, based upon neural data and a generalized optimality theory, incorporates a collapsing threshold (Drugowitsch, Moreno-Bote, Churchland, Shadlen, & Pouget, 2012; Ratcliff, Smith, Brown, & McKoon, 2016). In this model, the standard threshold parameter a is treated as non-constant, as "collapsing" over time (e.g., according to the functional form $a = \gamma e^{-\delta t}$ for appropriate parameter choices). This allows the diffusion model to capture the neurological concept of urgency-gating, where time to act is penalized independently of information evidence (Cisek, Puskas, & El-Murr, 2009; Thura, Beauregard-Racine, Fradet, & Cisek, 2012). This "cost to act" is meant to deal with an important issue with the diffusion model (Drugowitsch, Moreno-Bote, Churchland, Shadlen, & Pouget, 2012), that the threshold is set independently of the evidence quality. In situations where the evidence quality is low, the time to hit the threshold requires the addition of a noise term to the instantaneous information, e.g., $I = \sum v_t + \epsilon$, which allows the cumulative evidence I to pass the threshold. For instance, if v_t is close to 0, decisions are at chance and collecting more evidence does not substantially improve decisions. This makes the resulting decision both take a long time and be lower performance even when the added time cannot reasonably improve performance.

While urgency gating can resolve this issue, it could be better to track information reliability, and use this to determine decision time. Srivastava, Johannes, Schrater, and Vul (2016) apply a similar notion to information reliability, *predictive volatility*, in an economic paradigm with probabilistic and unknown payoffs, to predict when participants will stop sampling information to make a choice from a set of bets. Predictive volatility tracks whether an estimation process has not saturated, and therefore whether to continue sampling information. Here we extend this idea, using a general notion of confidence as information reliability, a metacognitive variable, that participants track to determine when to stop sampling.

4.1.2 Current study

General purpose decision making models require more flexibility than the standard drift diffusion model in that many decisions have variable information rates, uncertain information rates, or even information stored in memory rather than as part of the stimulus. One strategy the brain may use to handle more complex decisions is to take a hierarchical approach, whereby incoming information is monitored by processes that can keep track of how much information was available during a decision. Such monitoring complements information integration and can provide additional information about the confidence associated with a decision provided by the monitoring process.

Consider a decision problem where time-varying stream of stimulus information $X_{1:t} = \{x_1, \cdots, x_t\}$ arrives with variable reliability about a decision y representing a judgment (e.g., leftwards or rightwards movement). Normally, we model optimal decision making as recursively computing $p(y|X_{1:t})$, making a decision when the probability exceeds a criterion for desired accuracy (e.g., 95% correct). The standard diffusion model treats all information sources as reducible to an overall decision criteria v_t, and then relies on noisy accumulation to a fixed threshold a. However, what happens when the stimulus information stops before the desired threshold is reached? Or when the reliability of $X_{1:t}$ has changed? In these circumstances a meta-level monitoring of the information that can predict the quality and quantity of information expected is essential.

Meta-cognitive tracking of information reliability

Quality of information can be reliably predicted under many circumstances. The input stream may contain contextual information about the reliability, including contrast, dot density, visual factors or scene statistics. In addition, when information reliability is set by the brain through attentional gain changes or resource allocation (Denison, Adler, Carrasco, & Ma, 2018), these controls could be used to generate a kind of "efference copy" forecast of the effects of allocation decisions on decision reliability. In either case we can view the brain as having access to a meta-level variable that acts on the underlying decision process by a coordinating process of 1) forecasting information quality and 2) adjusting decision strategies to accommodate time varying information.

Let z_t represent a reliability forecast variable that is either inferred from contextual information or set by the brain to account for resource allocation, such as attentional gain Carrasco (2011). This variable now informs the probability estimate of the decision variable by parameterizing the incoming data likelihood. Using Bayes' theorem and a log transform, we can do the following:

$$\log p(y|X_{1:t}, z_{1:t}) = \log \frac{p(x_t|y, z_t)p(y|X_{1:t-1}, z_{1:t-1})}{\sum_y p(x_t|y, z_t)p(y|X_{1:t-1}, z_{1:t-1})} \tag{4.1}$$

$$= -\log Z_t + \log p(x_t|y, z_t) + \log p(y|X_{1:t-1}, z_{1:t-1}) \tag{4.2}$$

This is the standard update equation for evidence in most sequential models (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006), including an evidence term dependent on z_t. The purpose of z_t then, is to provide a cue to the reliability of the incoming data, so it can be considered as a cue for the spread σ_t (e.g., variance, entropy) of the likelihood function $p(x_t|y, z_t)$. This can be seen if we split the likelihood through a factorization:

$$p(x_t|y, z_t) = \int_{\sigma_t} p(x_t|y, \sigma_t)p(\sigma_t|z_t)d\sigma_t$$

meaning that z_t is predictive of σ_t. Using this interpretation of z_t and the standard interpretation of the recursive update equation as Wald (1945) sequential likelihood ratio test, we can rewrite the likelihood ratio using this forecast. Recalling that the entropy of a Gaussian random variable is $\log(\|\Sigma\|)$ (where Σ is the variance), we have the conditional entropy $H(x_t|y) \approx \log(\sigma_t^2)$ which means that we can forecast the information content from z_t through:

$$R_t(z_t) = E_{p(\sigma_t|z_t)}\left[H(x_t|y)\right]$$

where R_t is the expected value of the entropy, given the forecast of σ_t by z_t. Similarly, we can rewrite R_t in terms of the log likelihood ratio, given that the likelihood ratio is a simple function of the variance for Gaussian (e.g., $f(\sigma_1^2/\sigma_0^2)$ for hypotheses 1 and 0):

$$\hat{I}_t(z_t) = E_{p(\sigma_t|z_t)}\left[\log p(x_t|y=1,\sigma_t) - \log p(x_t|y=0,\sigma_t)\right]$$

This is the expected information gained at each timestep. In other words, $\hat{I}_t$ is a forecast of the information gained based on z_t. We can therefore represent the impact of variable reliability on information available by scaling the externally presented information, J_t, by its current reliability due to the forecast. Let w_t represent the amount of information that is actually available to the decision after transduction and processing. Then,

$$\sum w_t J_t = \hat{I}_t(z_t)$$

is the time-varying information accumulated.

By tracking reliability through z_t, we propose the brain also has a noisy estimate of the information available: $\hat{w}_t = \hat{I}_t(z_t)/J_t$. Notice that $\sum \hat{w}_t$ forms a complementary representation of the information available at the decision, which is not the same as the actual integrated evidence but instead represents a principled prediction based on the forecasted reliability of the evidence. We propose the brain uses this kind of higher order reliability forecast to both adaptively set decision thresholds and to provide a meta-level estimate of decision confidence.

Using a meta-cognitive confidence for adaptive decision thresholding

Given this meta-cognitive information the brain has access to, that of the reliability of the integrated information used in the decision, how should the decision be made? As previously mentioned, a threshold could be set based upon some previously determined accuracy criteria. However, if the brain has access to a reliability forecast, then it could stop sampling once it knows no new information can be acquired, as Srivastava, Johannes, Schrater, and Vul (2016) suggest.

We allow for an adaptive threshold setting to determine stopping. Let $C_t = \sum \hat{w}_t$ be the cumulative information available, and $\Delta C = [C_t - C_{t-1}]$. Information sampling is terminated when *either*:

1. $\Delta C < \epsilon$
2. $C_t > \theta(p_e)$

where $\theta(p_e)$ is a threshold set by some error rate p_e, and ϵ specifies the required saturation level. In other words, when the information available saturates, stop sampling and decide based on the likelihood ratio. Otherwise, stop when a threshold is reached.

These rules extend the standard stopping approach, accumulating information to some threshold, to include situations where information saturates or quality changes dynamically. An overview of this view is in figure 4.1, where the standard information integration is modified with a monitor that can set or observe the information quality, and adjust the threshold rule based upon that quality.

Importantly, decision accuracy can also be forecasted using C_t. This is because the information at response scales the probability of being correct (Wagenmakers, Van Der Maas, & Grasman, 2007; Palmer, Huk, & Shadlen, 2005). Hence, we can use our forecast: $p(correct|\ information) \propto logistic(\hat{I}_t(z_t)) \propto logistic(C_t)$ (where $logistic(*)$ refers to the logistic function). While the exact equality may not hold, it provides a prediction of performance accuracy before the feedback is received (up to some monotonic scaling). This allows C_t to be a measure of confidence that both predicts performance and impacts choice decisions.

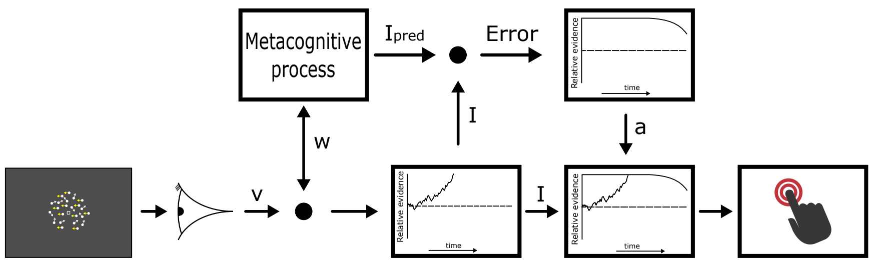

Figure 4.1: Conceptual overview of how information is integrated in a decision making process. Information comes in through sensory systems and decision-relevant information v is filtered through attentional processes. Information is accumulated over time (I) and compared against some threshold criterion a (e.g., based on error rates). Once this threshold is met a decision is made, producing overt action. While this process is occurring, a metacognitive process is monitoring or setting the attentional filter (w), producing a predicted information accumulation I_{pred}, and using the error in this prediction to adjust criterion levels (i.e., threshold boundary a) for decision making.

Analysis overview

We now update the diffusion model, based on our understanding of a metacognitive monitor that tracks reliability. In the standard diffusion model, the agent received information at each timestep, v_τ, that represents relative information in the stimuli with respect to one of the two hypotheses (e.g., $v_\tau > 0$ for the log likelihood of rightward motion). v_τ is the *drift rate* in standard decision models, and is combined with an error term ϵ_τ and integrated over time steps to produce the cumulative information I_τ. That is:

$$I_\tau = \sum_{\tau=0}^{T} v_\tau + \epsilon_\tau$$

However, we allow this information to be impacted by a weight factor at each timestep w_τ, e.g.,:

$$I_\tau = \sum_{\tau=0}^{T} w_\tau (v_\tau + \epsilon_\tau)$$

This represents the relative weighting as mentioned above. Here $I_\tau \approx \hat{I}_t(z_t)$.

Decisions are then based on the value of the cumulative information I_τ compared with a decision threshold a_τ, that is if $|I_\tau| > a_\tau$ then stop accumulating information and respond. Note that in the standard diffusion model, a_τ is constant (i.e., a constant decision threshold). Here we instead use a hazard analyses approach to investigate how the decision threshold is adapted based on the information available w_t. Hazard functions are a way of describing the instantaneous likelihood of an event such as a response (Luce, 1986; Ratcliff & Van Dongen, 2011). The hazard function $\lambda(t)$ is defined as:

$$\begin{aligned}\lambda(t) &= \lim_{dt \to 0} \frac{P(t \leq T < t + dt | T \geq t)}{dt} \\ &= \frac{f(t)}{S(t)}\end{aligned}$$

for probability density $f(t)$ and survival function $S(t) = P(T > t)$. In other words, the instantaneous chance of an event occurring at time t. We perform nonparametric fits of our response time data to characterize how the response curve is impacted by confidence.

In order to identify information integration weights w_t, we adapt a systems identification method from control theory. A standard way of characterizing a system is via the impulse response, that is, observing the output of a system after it experiences a brief input signal (i.e., the “impulse”) (Jagacinski & Flach, 2003; Shadmehr & Mussa-Ivaldi, 2012). We used a standard perceptual choice task, the random dot motion paradigm with two alternative

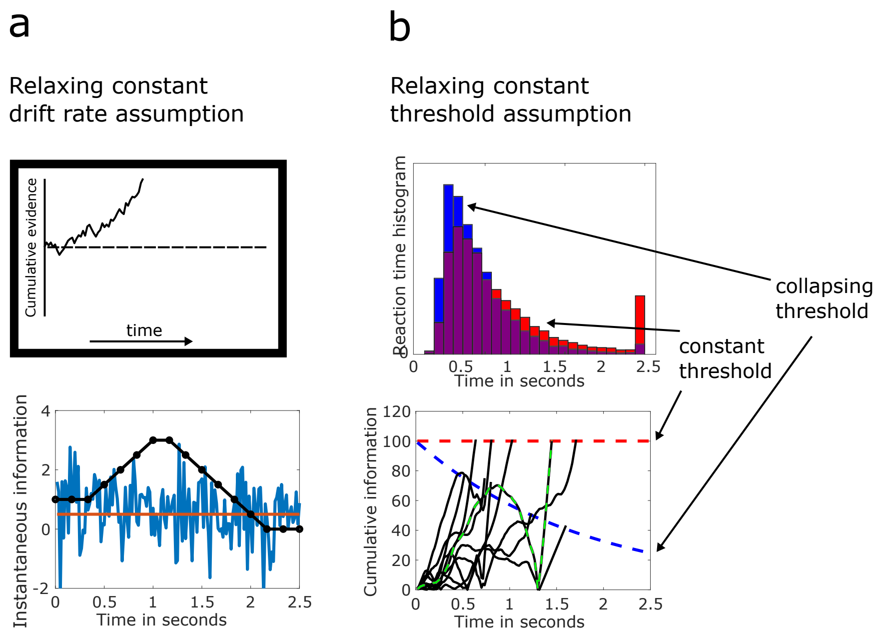

Figure 4.2: Conceptual illustration of generalization of the drift diffusion model. (a) We allow for time-varying weights to the instantaneous information state . While externally presented information v_τ is constant with noise, weights w_t are added to them to change the impact the information has on integration. The resulting instantaneous information is then summed to cumulative information I. (b) Here we allow for time varying thresholds, which produce distinct response time histograms. We can then identify this system by using an impulse response method.

forced choices (Newsome, Britten, & Movshon, 1989), with a thresholded motion coherence rate that functions as a standard for performance comparison. Coherence determines the signal intensity, so to simulate an impulse we added “bursts” of information, that is a brief but sharp increase in the motion coherence, to the stimuli at different points. We then applied logistic regression on choice accuracy to identify weight on information, and how weights vary based on confidence. The information I_τ above should reflect the probability of correct at choice time RT (Wagenmakers, Van Der Maas, & Grasman, 2007; Palmer, Huk, & Shadlen, 2005), that is:

$$p(correct|\,information\ at\ choice) \approx logistic(I(\bar{w}, RT))$$
$$= logistic(\bar{v}_\tau(RT) - \bar{w}_\tau)$$

Therefore, we can perform a logistic regression on a design matrix associated with when the burst (impulse) for a given trial occurs, and use this to find the unobserved $\bar{w}_\tau$.

We find that confidence reflects trial-by-trial information gain. It impacts response time by either delaying decisions till when information is useful, or responding early when information saturates. This suggests our internal awareness of low information gain translates into a longer integration time. Confidence then represents internal sensing of information gain, and adapts integration time to help offset the low gain.

4.2 Method

4.2.1 Participants

We collected data from 75 participants. All participants provided written informed consent and all procedures were approved by the ethics committee of the Institutional Review Board of the University of Minnesota. Participants were compensated either through cash or University course extra credit. They had normal or corrected-to-normal vision and hearing.

4.2.2 Procedure

Participants performed a variation of the random dot motion (RDM) paradigm (Newsome, Britten, & Movshon, 1989). Stimuli were presented through the use of the Psychophysics Toolbox in MATLAB (Brainard, 1997), based on custom MATLAB code originally developed G.M. Boynton at the University of Washington (`http://courses.washington.edu/matlab1/matlab/`). Participants were placed in front of a computer monitor and had access to a mouse for behavioral responses. Each participant underwent a thresholding procedure to set task difficulty such that individualized accuracy was 75 percent at experiment beginning. Through several rounds of administration, this additionally served to familiarize participants with the task. The main experiment ran for 400 trials per participant, in which their coherence level was constant and set based on the thresholding.

Participants saw a cloud of white dots moving on a black background. The cloud took up 8 degrees of visual angle. Briefly, some subset of these dots move coherently left or right while the rest move in random directions. Each dot in the cloud appears, moves a short distance, then disappears. It is the participants' task to gauge which direction they perceive the coherent subset of dots to be moving and to make a corresponding mouse movement as quickly and as accurately as possible.

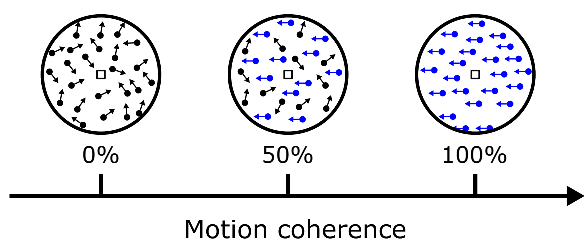

Figure 4.3: Example of coherence

Participants responded with both their decision and confidence simultaneously (see figure 4.4), using a mouse. This allowed us to avoid biases associated with post-decision confidence rating (Navajas, Bahrami, & Latham, 2016). A mouse allowed participants to move a bar to either the left or right to indicate rightward or leftward motion, by moving left they indicated leftward coherence, and right with rightward coherence. Moving the mouse either direction caused yellow bars to appear below the stimuli in the direction of their mouse movement, from one bar for less mouse movement to five bars. Their confidence was reported by the level of bars they revealed when they made right or left motion; if they revealed one bar they rated their confidence as 1, while if they revealed all 5 bars they rated their confidence as 5. Upon clicking the mouse they made their decision on both the direction of coherence and their confidence. Participants were informed of this and indicated their understanding before engaging in thresholding. Participants responded this way (i.e., with confidence and direction) during both thresholding and the main experiment, in order to become accustomed to the experimental design. After the experiment participants were debriefed.

If participants responded within the two and a half (2.5) seconds the dots disappeared. If they did not respond within the 2.5 second time interval the dots disappeared, the trial was marked as incorrect and the experiment continued. In either case participants could start the next trial by moving their mouse to the center of the screen (indicated by using the mouse to move a small square into a larger rectangle) and clicking the mouse, which started the next trial. Participants were therefore self-paced in the experiment.

In order to perform signal identification, on the main experimental trials we included a burst of information that is randomly selected at different points through the trial, either at frame 18, 36, 54, 72, or 90, (at 60 frames per second, corresponding to 0.3, 0.6, 0.9, 1.2, or 1.5 seconds), that lasts for 9 frames (0.15 seconds). On those frames, the coherence was increased, to be 4 times the coherence level the participant was thresholded at. The bursts are placed prior to and at 1.5 seconds, based on preliminary mean reaction times of 1.2 to 1.8 seconds for an individual. This burst of information was equally random across trials and participant with 1/7 probability. On 1/7 of the trials no burst was shown.

4.2.3 Data Analysis

A participant's data was removed if the participant did not use the full range of confidence ratings. In particular, if the standard deviation of their confidence rating was less than 0.5

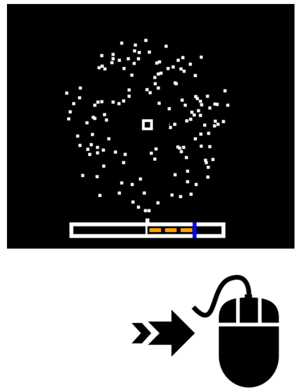

Figure 4.4: Participants indicate their confidence and choice simultaneously by moving the mouse to either the left or right (to indicate either a left or right movement decision). Their current choice and confidence is indicated by the vertical blue bar below the stimuli. The farther the blue bar is from the center point, the more confidence they rate their decision. As they move the blue bar, horizontal yellow bars appear to provide integer valued ratings for the participants (one yellow bar indicates a confidence of 1, while 5 yellow bars indicate a confidence of 5). Participants clicked on the left mouse button to indicate their choice.

(i.e., they typically only used one rating) and their mean confidence was below 1, their data was not used. 19 participants were removed for this purpose, leaving 56 for the remaining analysis. Including the left-out participants for the analysis did not significantly change results, however we only report data on those 56 participants.

We performed a bootstrapped logistic regression using custom scripts in R, along with the `glm` function with binomial family of distributions and the logit linking options. The design matrix was created based on participants expected stimuli information (see example figure 4.5). Each trial has 150 frames of stimuli data, including 9 frames of burst data (i.e., higher coherence stimuli). We lumped by 9 frames (i.e., the length of the burst) to produce a design matrix of coherence values over time, that is, 16 coherence values over the trial. When participants responded, we removed the stimuli, and so treated the coherence value as 0. We regressed against the participant's accuracy score (1 or 0), with no response treated as a miss.

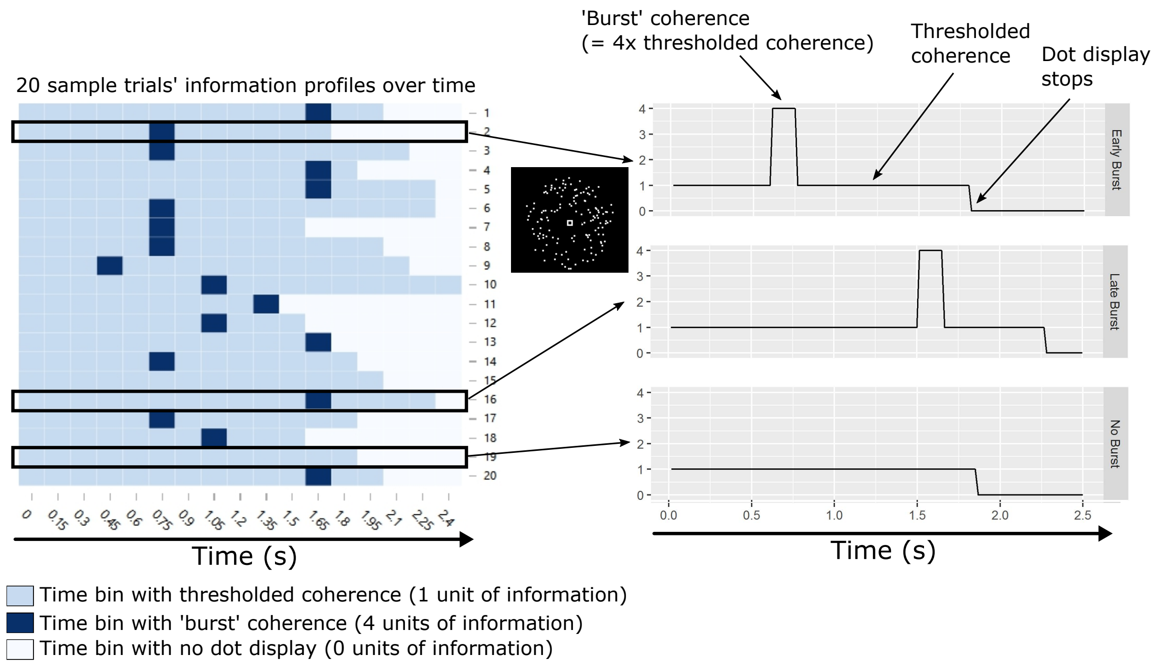

Figure 4.5: Example design matrix for 20 trials with three sample information-over-time profiles highlighted. Time is descritized over the x-axis for illustrative purposes. Approximate information available per unit time is highlighted either with color (left) or height of line (right). Note that the information presented here is based on expected levels of information; since the stimuli on a trial are random, the actual information available in the stimulus fluctuates. Also note that when the subject responds, we set the information level to 0 and the subject waits until the end of the 2.5 second trial before the next trial begins. Otherwise the information level is set at the individual's thresholded dot coherence value.

Since this design matrix is not full-rank, i.e., the columns are not linearly independent due to correlation in the covariates, we used a regularized regression (ridge regression), implemented via glmnet in R[2].

We also fit hazard functions to response time data, using a nonparametric Markov gamma process (Nieto-Barajas & Walker, 2002), a type of piece-wise hazard function fit that specifies a gamma prior process for the hazard functions, acting as a smoothing prior for our fits across continuous-time events. We used the BGPhazard R package[3] to produce hazard

[2]See https://web.stanford.edu/~hastie/glmnet/glmnet_alpha.html
[3]See https://cran.r-project.org/web/packages/BGPhazard/index.html

rates for participants' response time. Hazard functions were fit across participants, but separately fit within both confidence level and burst times (e.g., separate fit for confidence 1 and burst 18). The BGPhazard package fits a Bayesian nonparametric hazard rate via Gibbs sampling, and we specifically implemented the Markov gamma model using the function GaMRes. We specified 15 same-length partitioned intervals over the trial length, which determines the coarseness of our fit. This produces an estimate of the posterior hazard rate at each of these time intervals, for each burst and confidence score. The sampler used 3000 iterations, with both thinning and burn in of 600, which produced stable posterior chains based on diagnostic graphs and autocorrelation of the chains. We then treated the posterior samples as a histogram of the distribution over the hazard rates.

We then performed principal components analyses (PCA) on the sample hazard rates, in order to better investigate the impact burst timing and confidence had on response. We performed PCA across the average burst and confidence hazard rates, treating the the 15 time intervals as feature space to produce temporal components. PCA was performed using custom MATLAB code, producing components that represent principal hazard functions across time. We then projected, for each burst and confidence, the difference from the global average hazard rate onto the first two PCA components, which shows how the given burst and confidence score impacts the overall hazard rate.

4.3 Results

4.3.1 Basic visualization

On average, confidence has a negative relationship with subject reaction time (see Figure 4.6). Note here that a confidence of 0 represents a response with minimal movement, so while the mouse is to the left or right as indicated by the blue vertical bar, no yellow horizontal bars are graphically revealed. However, the relationship between mean response time and confidence is mediated by the timing of the extra burst information (see Figure 4.8. Note the "U-shaped" pattern; at low or high confidence, the change in response time is highly contingent on the burst timing, while within confidence values of 2-3 the response time change is much less impacted by burst-timing. In general, earlier burst times are associated with a decrease in mean response time, while later burst times have a moderate increase.

Confidence also appears to moderate the speed-accuracy trade-off (see Figure 4.9). Note that when confidence increases, accuracy also increases for earlier response times. This means that confidence increasing also increases overall performance, without impacting response time.

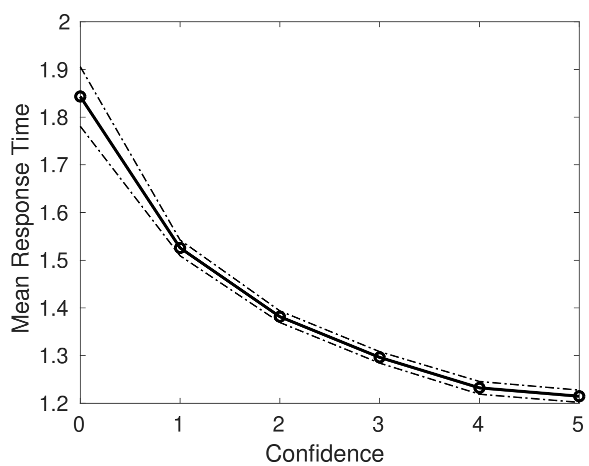

Figure 4.6: Mean response time for each confidence bin, along with 95% confidence intervals.

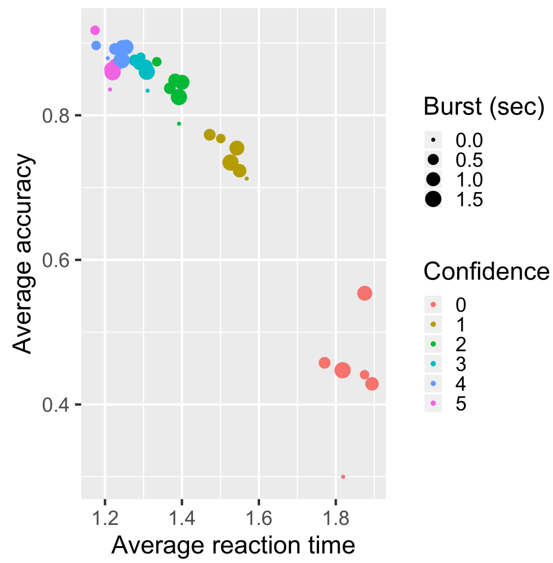

Figure 4.7: Average speed versus accuracy, grouped by confidence and burst time (seconds).

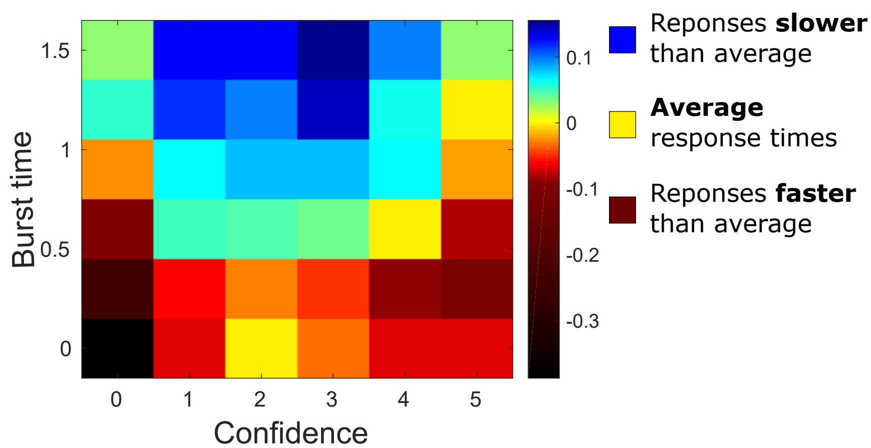

Figure 4.8: Color indicates a change in Mean RT as a function of burst time and confidence, with burst time indicated in seconds. Note that burst time takes frame values of either 18, 36, 54, 72, 90, or 0 (i.e., not shown) which correspond to second values of 0.3, 0.6, 0.9, 1.2, 1.5 or 0. Negative values indicate faster response, while positive are slower.

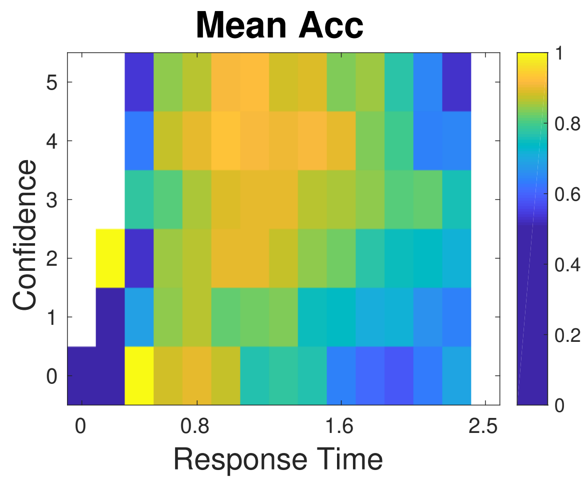

Figure 4.9: Color indicates mean accuracy for binned trials (based on response time and confidence). Bins which are white have no trials to sample from. Note as confidence increases, both mean accuracy and response times increases.

4.3.2 Logistic Regression

We performed a bootstraped logistic regression for each confidence level, and took the average regression weights as shown in Figure 4.10. Confidence bounds for the average weights were very small (all less than 10^{-10}), indicating significant weights. Note that the logistic regression weights are non-constant over time. All have an increased information weight to the midpoint of the timing period, that is, early and late information is down weighted while mid information is more predictive of accuracy. This interacts with confidence; in general, at higher confidence levels the information is more predictive of accuracy. As confidence increases, the information overall is more predictive of performance than at lower confidence levels.

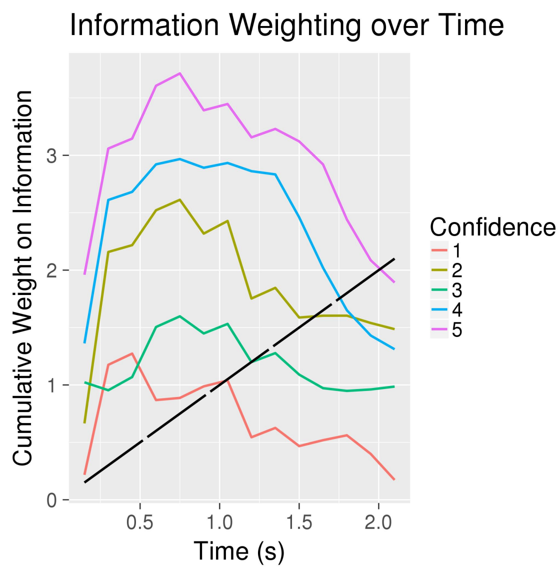

Figure 4.10: Plotted cumulative weights over time, from the logistic regression fits. The dashed black bar indicates a constant weight, which makes the cumulative weights linearly increasing. Accuracy was fit independently for each confidence level (as indicated by color). Confidence intervals for bootstrapped fits were very small (10^{-10}), indicating significant fits.

4.3.3 Hazard analysis

Resulting hazard rates are plotted in Figure 4.11, which were fit within both burst timing and confidence rating. Each line shows the hazard function for a given burst timing and confidence rating, indicating the instantaneous rate of response. Note the interaction between confidence and burst time on the hazard of responding. Later burst time generally results in a shifted hazard of response, producing later responses overall. However, the hazard of responding in general decreases as confidence decreases. This means confidence acts as a general gain, increasing the rate of responding across time (e.g., not just producing more early responses).

After performing PCA on the resulting hazard functions, we computed scores for each of the hazard function fits. Plotted in Figure 4.11 are the first and second component coefficient scores, for each burst and confidence hazard fit, along with illustrated plots below, indicating the impact that the components have on the resulting hazard function. This again demonstrates on average how a given burst timing and confidence rating impacts the hazard rate of responding. Note the interacting effect on response times; confidence acts as a general increase on the hazard, while burst produces a change on the shape (resulting in increasing hazard for later bursts, and an inverse U shape for earlier bursts).

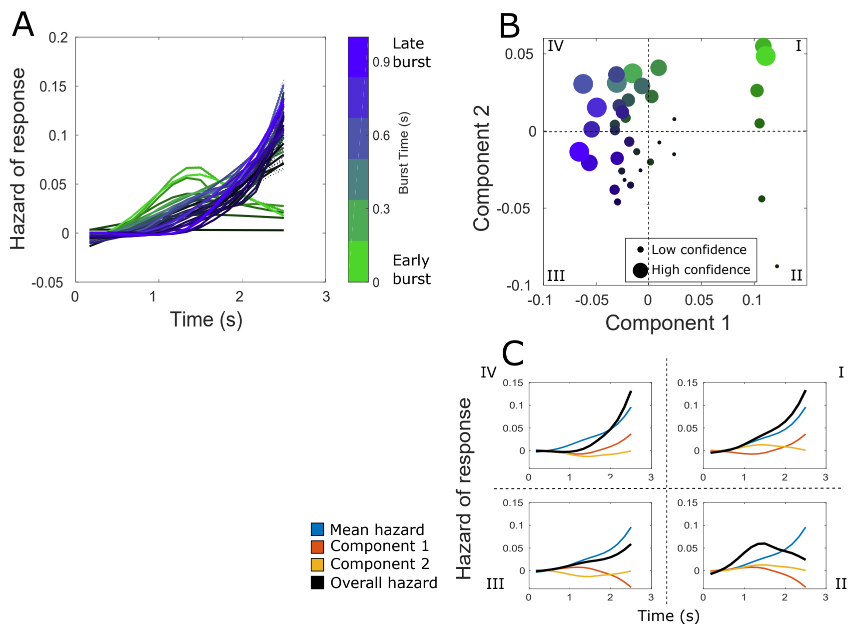

Figure 4.11: **(A)** Mean fit hazard functions, by both burst and confidence (36 fits in total, including unseen burst and 0 confidence). Burst timing is indicated by color (from green for earlier to blue for later, as a fraction of the overall trial length), while confidence is indicated by brightness (darker colors are lower confidence, while brighter are higher confidence). Time is in deciseconds (i.e., 250 deciseconds is 2.5 seconds). **(B)** Projected scores for each burst and confidence hazard function fit, on the first two PCA components. Burst timing is indicated by color (from green for earlier to blue for later, as a fraction of the overall trial length), while confidence is indicated by brightness and size (small, darker circles are lower confidence, while large, brighter circles are higher confidence). Component values are arbitrary, and simply indicate the relationship. **(C)** Quadrant figures demonstrate the impact the differing values each component has on the overall hazard function for each quadrant in the component plot. The mean hazard function (overall) is plotted in blue, while component 1 is red and component 2 in yellow. Note how the different quadrants produce shifts in the appropriate hazard functions (i.e., quadrant 1 and 2 have the same component 2 hazard, while shifting component 1 hazard). The resulting overall hazard function the quadrant has is in black, i.e., the result of adding the two components plus mean together. Note that confidence increases moving from quadrant 3 to 1, while burst increases from quadrant 2 to 4, allowing an easy comparison of the overall hazard.

4.4 Discussion

We investigated confidence by using impulse responses with a random dot motion task and collecting participant confidence judgments at decision time. We used the impulse response to perform logistic regression to extract weights on external information. We also performed hazard analysis to determine how human response time is impacted by both confidence and the bursts of information. We find that confidence reflects trial-by-trial information gain, which impacts both performance and response time. Confidence impacts response by allowing the decision process to delay when information is useful, or respond earlier when it is not. This suggests our internal awareness of low information gain translates into a longer integration time. Confidence acts as an internal sensing of information reliability, and adapts integration time to help offset the low gain. Confidence can achieve this if it acts as a meta-cognitive tracker, dynamically predicting performance and setting response.

Our findings fit with research indicating that confidence is a metacognitive variable that tracks information reliability (Yeung & Summerfield, 2012; Mamassian, 2016). Our results also align with confidence being positively related with attention gain (Denison, Adler, Carrasco, & Ma, 2018), a key source of information reliability for decision tasks. Interestingly, our results also align with neural evidence which shows that direct stimulation of sensory areas does not impact confidence reports (Rahnev, Maniscalco, Luber, Lau, & Lisanby, 2012; Fetsch, Kiani, Newsome, & Shadlen, 2014; Peters et al., 2017). If confidence reflects a forecasted reliability variable, rather than an instantaneous estimate, than current sensory information should be only distally related to confidence.

A possible limitation of our impulse response method is that while such a method works for analyzing linear time-invariant systems, which seems to well approximate standard perceptual tasks (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006), people are non-linear systems along longer time lags. For instance, participants might change their strategy of response if they are explicitly aware of the bursts. However few of our participants reported noticing the bursts when interviewed during post-experiment debriefing, suggesting our results are not due to explicit strategies.

Future work could look at more formal derivations of the diffusion model that utilize confidence as a reliability variable (Drugowitsch, Moreno-Bote, Churchland, Shadlen, & Pouget, 2012; Srivastava, Johannes, Schrater, & Vul, 2016). Such work could then allow for formal model comparison of alternative versions of confidence in decision tasks (Adler & Ma, 2018). Extending to multi-choice problems would also allow future work to investigate changes of mind as impacted by in the moment confidence. Presumably confidence will track changes of mind, as low confidence will indicate unreliable information and prompt more noisy decisions.

Interlude

Our priority queue architecture described in Chapter 2 requires computing a priority value and using this to decide when to quit a task. We now expand on the resulting impact this scheduling should have on people's time use and task switching, by making connections to optimal foraging theory. By framing time allocation as task foraging, we derive results that make practical and qualitative implications on human time use.

Chapter 5

Time Allocation as Task Foraging

5.1 Introduction

How long should we work on writing a paper before checking email, going to the bathroom or getting a snack? More generally, what determines how we allocate our time among these possible tasks? Our brains constantly make decisions about which activities to engage in, despite these activities subserving competing needs, having variable urgency, and having completion times that differ by orders of magnitude. Moreover, long term tasks such as completing a journal paper, training to run a marathon, or beating the high score of a video game require repeated engagement and intermittent scheduling. Successfully allocating time across diverse needs, goals and time scales is a critically important but challenging computational problem.

Unsurprisingly, people are often unsuccessful at maintaining engagement in long term tasks and often quit or churn from a task, either temporarily or permanently (see Figure 5.1). Sometimes this churning is caused by external events (e.g., a ringing phone) or internal events (e.g., hunger or fatigue). However the churn-causing event is often inexplicable; an individual may decide to stop working and check their email for no apparent reason. Here we develop the idea that spontaneous quitting is a natural consequence of a rational, subconscious probabilistic process for time allocation.

In psychology, spontaneous quitting (also referred to as *self-interruptions* or *sponta-*

neous task switching) is usually framed in terms of cognitive control and self-regulation theories of motivation, such as those put forth by Miyake et al. (2000), Carver and Scheier (1998) and Hofmann, Schmeichel, and Baddeley (2012). An individual quitting a particular task is often identified as lacking certain executive abilities, energy, or "willpower" (Baumeister & Heatherton, 1996). Similarly, impulsive behavior is treated as a deficit in a regulatory mechanism that keeps the individual "on task." These views often ignore the frequent ecological benefits of switching tasks due to new opportunities, changes in need states or diminishing returns (Wrosch, Scheier, Miller, Schulz, & Carver, 2003a). In fact, keeping on task when the task devalues, progresses much slower than expectation, or with doubtful completion is also viewed as irrational, typically referred to as a *sunk cost effect*. Moreover, dynamic changes in the priority of our intrinsic needs (e.g., hunger or thirst), cognitive fatigue and the availability of rewarding short-term options provide a rich set of rationales for reallocating our time away from long-term activities with uncertain outcomes. Blaming self-control for quitting early and stubbornness for quitting late are both suspect judgments uninformed by the complex factors our brains use to manage time allocation.

A similar moral perspective is notably absent in analyses of time allocation in animals. Instead, foraging theory (Stephens & Krebs, 1986; Stephens, Brown, & Ydenberg, 2007) provides a rich theoretical framework for predicting how animals allocate time for food search or mate selection as a rational, though possibly constrained, process that weighs costs and benefits of alternatives. From a foraging perspective, a decision to quit can be thought of as a rational time allocation choice given the values of the current task and its alternatives. However, scheduling arbitrary tasks is more complex than foraging — the dynamic value of a task involves integrating a richer array of kinds of costs, benefits and uncertainties than the accumulation of food that standard foraging theory provides.

Developing a theory of task scheduling requires integrating both external and internal

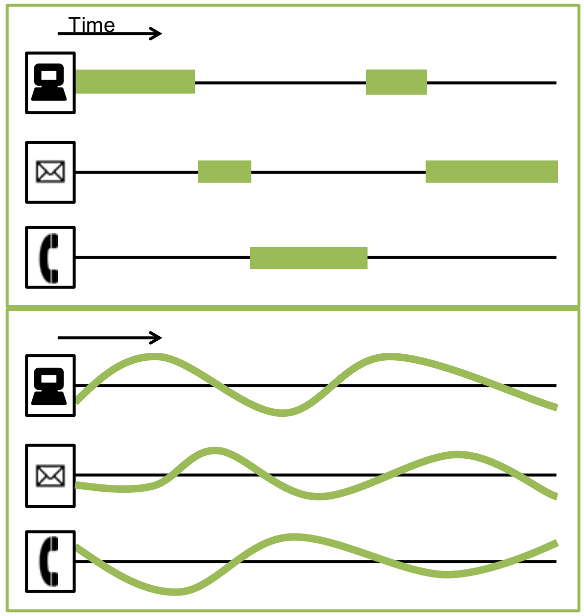

Figure 5.1: The phenomenon of spontaneous task switching. (upper) This Gantt chart shows a person's time spent on a task. Each horizontal line indicates a distinct, potential task, while the thick green bars indicate the current task at any particular time. In this example the three tasks might be writing a paper, checking work emails, and talking on the phone. (lower) The value of each available task will spontaneously fluctuate over time. This could be due to internal factors, such as boredom or fatigue, or external, such as availability or deadlines.

factors affecting time allocation into a common framework. While this scheduling problem forms a core problem of motivation theory, current motivation theories do not directly address scheduling from a time allocation perspective (Niv, Joel, & Dayan, 2006; Berridge, 2004; Carver & Scheier, 1998). In particular, while most motivational theories highlight relevant factors that influence time-in-task, they do not provide a way of combining these factors into a theory that can predict time allocation. We combine ideas from motivation and optimal foraging theory (Stephens & Krebs, 1986; Clark & Mangel, 2000) to build a novel quantitative theory of what makes people want to persist in or quit an activity. By treating time allocation as a control problem, we can view a person's emergent task-switching behavior as the result of rationally solving this scheduling problem.

5.2 Background on spontaneous task switching

Most insights and evidence pertaining to time allocation in humans come from a set of cognitive control and self-regulation paradigms that are strongly controlled and yield only indirect access to spontaneous time switching processes. These paradigms place strong external demands on subjects' time allocation while varying external pressures, and provide critical evidence for factors that influence and derail goal-directed time allocation.

Below we review previous research on spontaneous quitting and self-interruptions, along with effects that can impact an individual's spontaneous task switching. These effects are often qualitative in nature, and are necessary for a theory of task switching to explain. As noted by Gazzaley and Rosen (2016), many of these self-interrupting effects might be explicable through an understanding of foraging theory. Our work takes a similar view, but formally extends optimal foraging theory and works through those implications, taking seriously the difficulty of determining task priority.

5.2.1 Forced interruptions, multitasking and spontaneous task switching

Cognitive research on task switching generally focuses on the impact of forced interruptions and externally-induced time constraints. The interest here has more to do with how cognitive limitations might impact multitasking and vice-versa (Wang, Irwin, Cooper, & Srivastava, 2015). Humans have fundamental limitations in multitasking, in particular in scenarios where one goal, such as work, takes precedence against distracting tasks (Marulanda-Carter & Jackson, 2012). Task switching as an experimental paradigm is used to assess how well people can switch between different goals or policies, in an effort to understand the underlying mechanisms that are involved in this switch. Subjects are trained on multiple simple tasks and then switch between them (Monsell, 2003), either cued by a stimulus cued or allowed to switch voluntarily (Arrington & Logan, 2004). In either case, switching is experimentally *directed* rather than spontaneous. Typical results for directed switching show that a subject's performance suffers on switch trials in terms of reaction time and accuracy (Monsell, 2003). Some studies indicate that subjects are aware of cognitive costs associated with directed switching, and that these costs make switching aversive (Kool, McGuire, Rosen, & Botvinick, 2010). However despite this, people still will engage in spontaneous switching in the face of a distracting environment (Marulanda-Carter & Jackson, 2012).

Spontaneous task switching is generally unexplored in the task-switching literature. Primarily this is because task switching focuses on consequences of sudden goal switching rather than why one might switch goals (Hofmann, Schmeichel, & Baddeley, 2012). An exception is Kessler, Shencar, and Meiran (2009), who developed a spontaneous task switching paradigm that showed subjects freely switching tasks even when incurring a performance cost. Importantly, the authors note that spontaneous switching had not been previously experimentally verified, despite its widespread commonsense acceptance. They similarly note that contemporary theories of executive control cannot account for such phenomenon, which is understandable given the historical separation between executive function and self-regulation research (Hofmann, Schmeichel, & Baddeley, 2012).

5.2.2 Self-interruptions and impulsivity

While humans have difficulty in multitasking due to external interruptions, much of daily task switching is due to self-interruptions (Jin & Dabbish, 2009), often theoretically explained via limitations in self-regulation or self-control. Impulsivity is "acting without thinking," either in terms of choosing a short-term reward over preferred long-term rewards or making decisions habitually as opposed to deliberately. While increasing the rewarding value of a task can incentivize sticking, the presence of off-tasks rewards can incentivize quitting (Mischel, Ebbesen, & Zeiss, 1972). Self-control, in contrast, is when a human changes their own behavior explicitly in order to override the default, habitual, or stimulus driven processes that would otherwise control behavior (Carver & Scheier, 1998). Self-control and self-regulation are seen as opposing impulsive behavior, and failure in those mechanisms has been considered a candidate cause of much of our own unwanted behavior: e.g., drug abuse, overeating, gambling addiction, and violence (Baumeister & Heatherton, 1996). Real world problems and cognitive deficits are then attributed to a tendency towards failures of self-control (e.g., Greenberg and Waldman, 1993). People will pragmatically adjust to this limitation by reducing temptations for alternatives tasks, as we see in computer or phone applications that block unwanted websites.

A prototypical example of self-control limitations is the marshmallow experiment (Mischel, Ebbesen, & Zeiss, 1972). The experimenter places a marshmallow on a table in front of the child and tells them that if they wait until the experimenter returns, they will be given a second marshmallow. Generally the experimenter is gone for only 15 minutes, with the amount of time the child is willing to wait being recorded. The ability of the child to wait longer for the larger reward (i.e., 2 marshmallows) is often attributed to them having better self-control, having been correlated with long-term success in many real-world outcomes such as SAT scores (e.g., Mischel, Shoda, and Rodriguez (1989), but see Watts, Duncan, and Quan (2018) for an update and reinterpretation).

By comparing subject behavior against perfect performance, these studies implicitly prescribe a set of behaviors; the correct option is to wait for the second marshmallow. Treating poor performance as a deficit assumes subjects share the experimenter's goals and knowledge. However, the goals and knowledge of subjects may diverge from experimenter expectation; participants may have different certainties about rewards or different goals in the task than both other participants and the experimenter. To demonstrate this, Kidd, Palmeri, and

Aslin (2013) performed the marshmallow task on a group of children, but first showed the experimenter being either unreliable or reliable on an initial task. Children first completed an art project where the experimenter promised them a larger set of art tools (e.g., crayons), however, for some of the children they did not provide the promised item. The authors found that those children in the reliable experimenter condition waited significantly longer for the larger reward (i.e., extra marshmallows) than those in the unreliable condition.

This result prompts reinterpretation of the relationship between performance in these delayed reward tasks and the real world predictions. Delaying reward only makes sense if the future rewards have a low uncertainty. An understanding of the natural problems that individuals may face, like poverty (Mani, Mullainathan, Shafir, & Zhao, 2013)), points out limitations of these existing frameworks. In particular, the ecological relevance of experiments that treat delayed gratification as normative is questionable (Fawcett, McNamara, & Houston, 2012). Delayed-discounting experimental paradigms have been shown in recent work to be a difficult problem for animals to learn due to the unnatural structure of the task (Blanchard et al., 2015; Carter, Pedersen, & McCullough, 2015).

A "good" solution in a natural environment can appear irrational in unnatural settings (Stephens & Anderson, 2001; Stephens, 2008; Fawcett et al., 2014). Critically, in a natural setting humans (and animals) have multiple tasks they must complete, implying trade-offs due to time being a critical resource to be managed across tasks. Therefore, we eschew the standard prescriptive treatment of self-control, instead treating time-allocation as a control theoretic problem that incorporates the subject's own goals and beliefs in lieu of experimenter expectations that may or may not be shared. While this can be placed within standard architectures for human motivation, it emphasizes a distinct control variable that is often ignored: *time*.

5.2.3 Causal factors that induce quitting and sticking

Previous research has collected a large set of factors that can impact task switching, which can either inducing quitting or sticking in a task. However, incorporating and explaining why these factors impact switching has proven difficult. These phenomena include both specific results in the experimental literature and common observations concerning issues in task completion. Note that this list of phenomena is not necessarily exclusive, simply prototypical and otherwise difficult to explain.

- *Distracting environment:* Difficulty of engaging in a main task due to distractors, where an aversive main task requires long commitment (Draheim, Hicks, & Engle, 2016; Mark, Iqbal, Czerwinski, & Johns, 2015; Dabbish, Mark, & González, 2011). During high cost but important to complete tasks, massive switching will occur; when studying for finals or finishing a dissertation, students will focus on cleaning their house first.

- *Interruptions and notifications:* Appearance or reminder of alternative task can cause immediate switching when main task requires continuous work (Iqbal & Bailey, 2008; Mark, Iqbal, Czerwinski, & Johns, 2015). This commonly occurs with "pop-up notifications" for email or instant messages.

- *Dual effect of task blocking:* Many "nanny apps" that block alternative tasks people can engage in (e.g., blocking Facebook during a writing task) have dual effects. Low completers (i.e., individuals who have an otherwise hard time completing tasks) benefit from difficult transitions, while high completers do not (Mark, Czerwinski, & Iqbal, 2018).

- *Gamification and accountability:* Gamification is the application of game design ideas to other non-game tasks to encourage engagement. Some instances of gamification can improve time on task, provided that points or accountability closely track task progress (Dickey & Meier, 2005; Hoffman & Nadelson, 2010). Similarly, added social accountability can improve time on tasks, for instance writing groups can keep people on task to write regularly.

- *Quitting near task completion:* Some tasks require completion to be satisfied, but people will quit early (Carver, 2003; Kidd, Palmeri, & Aslin, 2013). This can occur when alternative tasks are available and the main task has diminishing returns.

- *Progress sweet spot (dual effect of progress):* Both high and low progress (relative to expectations) can cause switching. Quick progress on a single task can cause switching if other high rewarding tasks exist., however, too low progress can also cause switching (Koo & Fishbach, 2012; Schmidt & Dolis, 2009).

- *Dual effects of deadlines:* Deadlines cause switching when far away (i.e., no work until deadline looms) but sticking when near. Uncertainty associated with task completion alters whether deadlines incentivizes working or switching behavior (Hartonen & Alava, 2013; Jarmolowicz, Hayashi, & Pipkin, 2010).

We come back to these phenomena in table 5.1, demonstrating how they can be explained through understanding the structural results of time allocation.

5.3 Modeling spontaneous task switching phenomena

The properties of task switching that produce time allocation are quite general. However, there are general structural requirements for time allocation to occur. Importantly, all of these spontaneous task switching phenomena occur in the presence of multiple goals that can be pursued, and where switching might have costs associated but is possible. We reconceptualize many of the phenomena by first reformulating task switching as the control problem of time allocation. Time allocation is a constant, if implicit, problem that people must solve, due to the simple fact that many tasks we engage in are mutually exclusive.

5.3.1 Mutually exclusive goals require time allocation

There are various structural assumptions we must make to model the task structure of standard task switching. Figure 5.1 indicates the standard structure of spontaneous task switching (Judd, 2015). Over time people will switch in and out of various tasks, sometimes returning to a given task or interleaving other tasks.

Allocating time across different tasks is necessary when the tasks are mutually exclusive, i.e., when they cannot both be performed at the same time. By choosing to engage with a particular task humans and animals are also choosing to not engage in alternative tasks. If multiple tasks have to be completed, time must be allocated to these different tasks based on the values of these tasks. Rats, for instance, must alternate between consummatory, grooming, and mating behaviors (Niv, Daw, Joel, & Dayan, 2007). In ethology, transitions between tasks naturally occur as animals must trade off between different goals (Fentress, 1983). Foraging behavior represents a prototypical instance of this, where an animal must choose to forage in different patches of food, forcing them to allocate their time.

Charnov (Charnov, 1976) showed that maximizing an animal's average rate of gain by foraging can be treated as an optimal stopping problem that focuses on when to leave a patch of food (Oaten, 1977). Charnov's key result, the Marginal value theorem, has been applied across animals (Stephens & Krebs, 1986), including human foraging behavior (Smith, Bettinger, & Bishop, 1983), and has even been used to model how people forage for information (Pirolli, 2007). In information foraging theory, the information content on a website is treated as analogous to a patch of food. A person's goal is then to maximize information gain for a topic by allocating time among websites. This theory has been extensively developed and produced an array of practical results (Pirolli, 2007). We show that this theory can be extended further, as a framework to understand general task scheduling.

Foraging theory has been extended to other animal task regulations such as mating behavior or predator avoidance (Mangel & Clark, 1986; Houston, Clark, McNamara, & Mangel, 1988), usually through a *fitness function* (Houston & McNamara, 2014) (i.e., a proper choice of value variables) that serve as currency. For example, foraging for food should increase an animal's energy budget, while avoiding predators should allow the animal to stay alive (both of which allow for future reproductive success). The impact of task selection on a human's evolutionary fitness is not directly observable in the same way. Typical human tasks (e.g., video games, academia, or scrapbooking) reduce energy and produce almost nothing tangible. One challenge in generalizing foraging theory to encompass task scheduling is in finding an appropriate value function, that is, understanding what should prioritize tasks.

5.3.2 Modeling the engaging properties of tasks

We want to understand what factors convey a task's priority or *urgency*. Psychologically, it represents an instantaneous rate of desired engagement for a task and should allow for a prioritization among all available tasks. We don't believe this urgency signal is a solely hedonic reward because difficult tasks often come with negative valences like frustration, which nevertheless increase task persistence (Carver, 2003; D'Mello & Graesser, 2009; Berridge, Robinson, & Aldridge, 2009). Instead we expect urgency to be a global computation that incorporates reward, costs, progress, relative mastery, availability and many other signals of the overall desirability performing a task *now*.

A suitable function for urgency defines what goals an agent has and therefore what behavior it is optimized for. We want to extend foraging theory to other tasks besides that of energy consumption, while maintaining the mathematical properties of energy consumption's fitness function. In particular this new urgency function should be an increasing function of time spent in the task, and it should include depletion or satiation (i.e., depression à la (Charnov, 1976)); without depletion, a human would stick with one task and never quit.

We must also incorporate tasks that will provide reward upon completion, but still have signals that indicate task progress. Rate of progress, in terms of expectations, has been shown by Carver (2003) to keep people working towards their goal, provided their progress is neither less than expected or more. While progress alone cannot strictly predict engagement (e.g., progress is susceptible to framing effects Koo and Fishbach, 2012), it is an important component of the decision to stick or quit and must be integrated.

Our urgency signal should incorporate uncertainty of task completion and progress, as both appear to be general indicators of whether individuals stick at a task. As indicated by the extended version of the marshmallow task (Kidd, Palmeri, & Aslin, 2013), increasing a child's belief of goal completion promoted sticking in the waiting task. Similar ideas have been developed by Wrosch, Scheier, Carver, and Schulz (2003b) showing that goal disengagement is good if the goal is unattainable.

Our theory is based upon the idea that urgency refers to the rate of relevant, valued events that that people accrue due to engagement in a task. While there are various events that people monitor, some represent a valued satisfaction of underlying goals or needs that prompt engagement. In order to emphasize the dissociation between task engagement and hedonic reward, as well as to emphasize the dynamic characteristic of a task's importance or utility, *urgency* will be used throughout this text to refer to the signal that drives time allocation, rather than reward, fitness, or utility function. Although we use the term urgency, our use corresponds to a purely endogenous assessment. Externally imposed urgency such as deadlines can create demands that cannot be fulfilled, which then can impact the resulting urgency of a task endogenously.

5.4 Time Allocation Theory

Consider the example of a researcher working on a manuscript. During the task, she may spontaneously switch from writing to something else that is salient, like reading email, paying a cost for transition in terms of writing, but possibly benefiting from meeting some other obligation. Like a standard decision problem, the author has choices, beliefs and a cost/benefit analysis, but she also must choose how much and which resources to devote to a task, with uncertainty about the likelihood of completion and the value of the task. Time allocation problems like these can be modeled as *Continuous Time Partially Observed Decision Processes (CT-POMDP)*. Continuous Markov Decision Processes were developed to model operations research problems, such as scheduling factory jobs. The general problem of allocating time and resources to different tasks can be considered an optimal scheduling problem (Bertsekas, Bertsekas, Bertsekas, & Bertsekas, 1995). Standard solutions to time allocation problems take the form of a priority index across tasks (Gittins, Glazebrook, & Weber, 2011); each task is assigned a priority score based on local information, and the current task with the highest priority is worked on.

We take advantage of a series of results to simplify the full CT-POMDP. First, the CT-POMDP can be approximated in discrete time using *uniformization* (Rao & Teh, 2013). By choosing an appropriate timescale, e.g., the speed of the fastest transition rate or via a sampling process, we produce a standard POMDP. A standard POMDP can also be represented as a belief MDP; by marginalizing over the uncertain states, we can frame the problem as a much simpler decision process over belief states. Finally, we show in Chapter 2 that we can abstract over within-task decisions and focus on selection of tasks. When our solution space can be decomposed over mutually exclusive goals, where goals can be represented as satisfiable constraints, within-task dynamics can be averaged out and the problem becomes one of selecting goals over time, rather than actions within goals. In Chapter 2 we show the conditions under which time allocation occurs following a POMDP, which results in the architecture in Figure 5.2, and a scheduling problem that appears as much more of a classical decision problem; when do you work on what tasks (goals) and for how long?

Our abstraction over within-task dynamics is possible provided we focus on a set of most relevant tasks over a time period. We can then focus on a set of relevant, short term, and potentially satisfiable tasks that can be completed. While in principle, it might be necessary to specify how a person trades off against every conceivable goal, not just eating and writing but also going to the moon, for instance, in this chapter, we limit our focus and proof to those most available tasks. This is conceptually similar to the availability heuristic in decision making (Tversky & Kahneman, 1973), where only those most available tasks within memory are actually selected for scheduling. This constitutes a set of potential "background" tasks that are actually scheduled; those in our priority queue in Figure 5.2.

While the current set of background tasks appears static in our formal derivation below, it only needs to change at a significantly slower time scale than the time allocated to the current task. This can be considered a "linearization" of sorts; as long as the background tasks change much slower than the current task completion, then we can expand the rate equations around the background rate to derive qualitative influences. Here, we focus on the structural implications of dealing with this scheduling problem, showing how they can

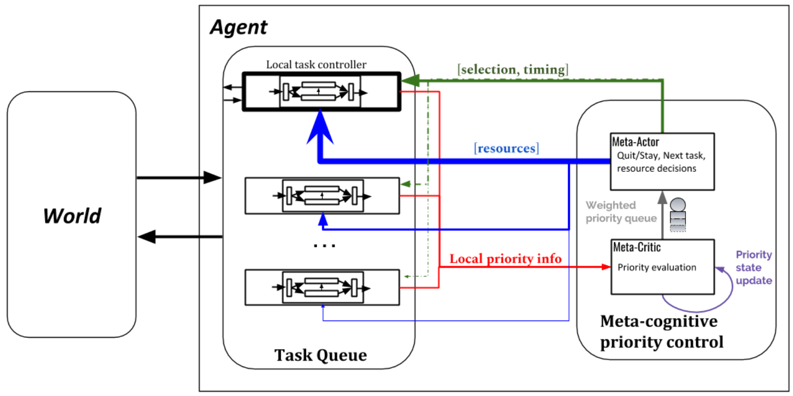

Figure 5.2: Engagement as the result of a metacognitive priority control system (originally in Chapter 2). Different local task actors determine which task goals are driving overt actions (within-task control). Outside is a meta-level controller that performs across-task control, assigning resources and activation to each task controller (e.g., by specifying goals for the task control loop). Local task actors send back local priority information from each task to the meta-critic. The meta-critic computes an overall priority score, integrating across longer time scales, while the actor computes priority scores across a salient task set ("queue"). We focus here on the problem of determining the priority score, however it is important to note that the meta-cognitive system produces action emissions that should be coordinated.

explain many of the phenomenon described above.

5.4.1 Optimal Time Allocation Theory for Task Foraging

In our constrained stopping problem, Figure 5.3 depicts the variables that should influence this decision of when to quit one task and switch. We construct an optimal control theoretic model of time allocation for *maximizing expected urgency rate*. We combine factors to compute a function for the urgency rate $\mathbb{u}(t_{w1}, t_{w2}, \dots, t_{wn})$ where t_{wj} is the time allocated to the j^{th} task[1]. The expected urgency U, is the average cumulative urgency experienced

[1] In RL, this would be the value function as a function of time

across tasks with time allocations t_{wj}, with expectation across all the factors listed above that influence urgency rate [2].

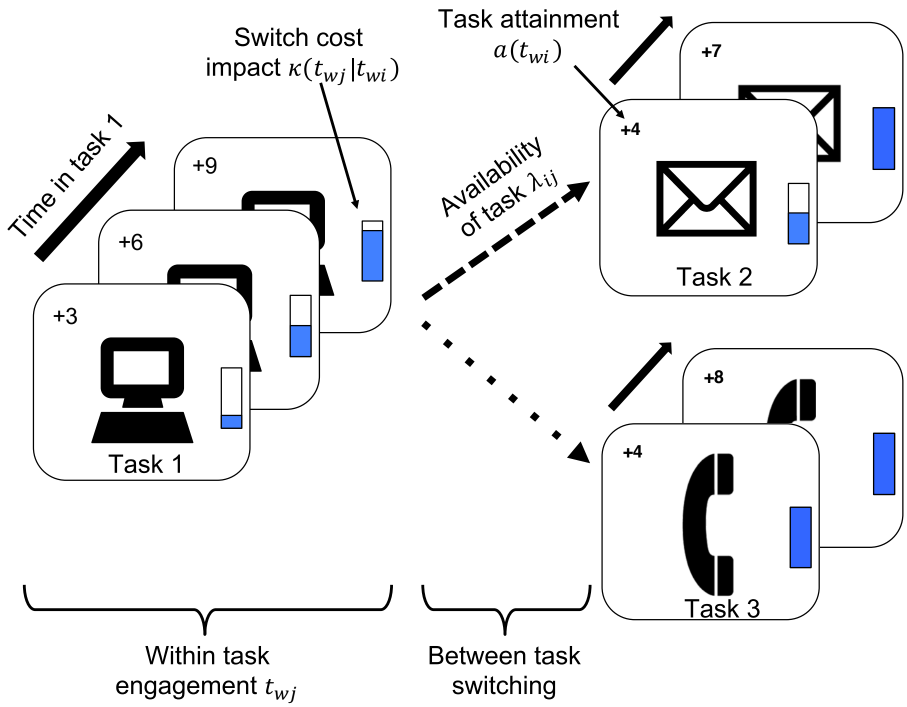

Figure 3. A set of available tasks: use computer, write email, and use phone. Properties of these tasks drive the decision process for determining how long to keep writing and when to switch. Panels indicated snapshots of time in a task, with the number in the top-left indicating the attainment for performing the task, and the bar on the right indicating switch cost.

The *Attainment* $a_j(t_{wj})$ within task j represents a forecast of the accrual of events that are rewarding, indicate task completion, or are internal events with their own intrinsic value. For any kind of event, attainment functions can be defined as: $a_j(t_{wj}) = P(N(event) \geq 0|t_{wj})$, that is, the cumulative probability of the number of these relevant events. Attainment is therefore a forecast of these satisfying events, which can be quite general as long as they

[2]This is a modification of Holin's Disk equation from (Stephens & Krebs, 1986); however, it can be drawn from an infinite horizon control problem (Bertsekas, Bertsekas, Bertsekas, & Bertsekas, 1995) or from renewal theory (Stephens & Charnov, 1982)

have some form of positive value [3]. We expand on relevant events later, but for now it's important to emphasize that they can be either external or internal events (e.g., cognitive or biophysical). Functionally, we assume attainment is a monotonic increasing function of the time on task, and also assume that $a_j(t_{wj})$ has a decreasing but non-negative derivative. This is equivalent to assuming that time on task decreases urgency, which prevents any one task from monopolizing time allocation.

The urgency of the task is then gated by whether the task can be performed from the user's current state (the task is available). The *Availability* of the task j, represented by a rate λ_{ij}, depends on what task you are leaving from. This captures the intuition that whatever task you are doing changes how available other tasks are. For example, when working with a text editor your email is highly available. However, email is less available while using your phone or eating lunch. This availability acts equivalently to a switching cost by discounting tasks that are more difficult (time/effort) to switch into.

Finally, we also incorporate *Switching Costs* between tasks. Moving from one task to another often produces a cognitive cost associated with that switch, which can produce a functional loss in progress relative to baseline. We model this conditional cost using $\kappa_i(t_{wj}|t_{wi})$, a matrix associated with the loss of attainment for task j in time t_{wj} after having spent t_{wi} time in task i. In general, this function is a probability (i.e., takes values between 0 and 1) and increases with t_{wj} and decreases with t_{wi}; it is the total fraction of events kept given the possible loss due to the time in the alternative task. Importantly, while we assume this for our proof, we demonstrate that it is unneeded to develop our results.

Overall, our assumptions for this model are:

1. The rate of availability of task type j given that you are in task i is a constant.

[3] These can be thought of as signals of need or goal satisfaction, as in Chapter 2.

2. The expected net gain of attainment in a task is non-decreasing with decreasing non-negative derivative (i.e., the attainment is quasi-concave).

3. The switch cost increases with the time in the task and decreases with time spent away from the task.

Considering these assumptions, we construct an expected attainment over the time allocated within tasks T_w and transitioning between them T_B:

$$A = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_{ij} T_B a_j(t_{wj}) \kappa_j(t_{wj}|t_{wi}) \tag{5.1}$$

This produces an expected urgency rate:

$$\mathbb{U} = \frac{A}{T_w + T_B} \tag{5.2}$$

Note that this function represents an expectation across beliefs about task availability, task completability, and urgency signal availability, and includes dependencies between tasks; it is not just the simple average of attainment. Despite the complexity we show (see supplemental section 5.9) that optimizing $\mathbb{U}$ produces a simple optimality condition, that is:

$$\frac{\partial \hat{A}_j}{\partial t_{wj}} = \mathbb{U}^* \tag{5.3}$$

meaning that the optimal time in a task is such that the instantaneous rate of gain in the task is equal to the average rate of gain at the optimal time allocations.

Equation 5.3 defines an optimal switching criteria: an agent should leave a task when the current instantaneous chance of urgency gain ($\partial \hat{A}_j / \partial t_{wj}$) is equal to the average rate of gains in the environment ($\mathbb{U}^*$). In other words, if the average rate elsewhere is better than you have now, leave the task.

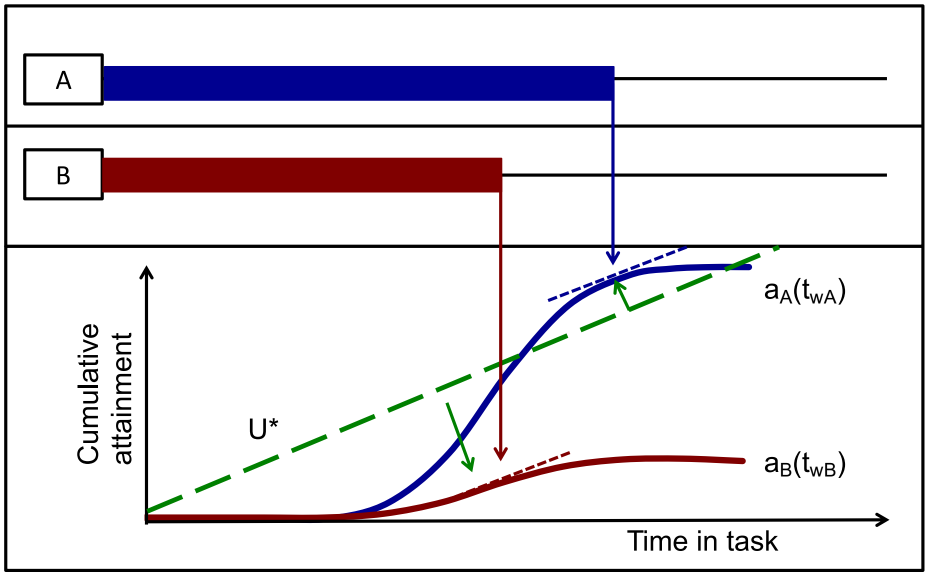

Figure 4. (Top) Optimal time in task as determined by marginal value theorem. (Bottom) The optimal switching point is when the rate of attainment drops below the average environment rate. This is visualized as the tangent lines whose slope is equal to the background attainment rate $\mathbb{U}^*$.

Taking equation 5.1 and removing the dependencies on task i from λ_{ij}, and removing κ_j, produces a result that is equivalent to the Marginal Value Theorem from optimal foraging theory. Equation 5.3 represents a generalization of the original result from Charnov (1976) with minimal modifications to capture variables important in most human tasks.

5.4.2 Structural Analysis of the Time Allocation Solution

Here we draw out the structural relationships between the different parameters of the model and their implications on time allocation, through the use of equation 5.3.

Graphical analysis

In foraging, optimal time allocation results are often presented graphically, as in Figure 5.4. The condition specified by Equation 5.3 means that the optimal switching point is simply the

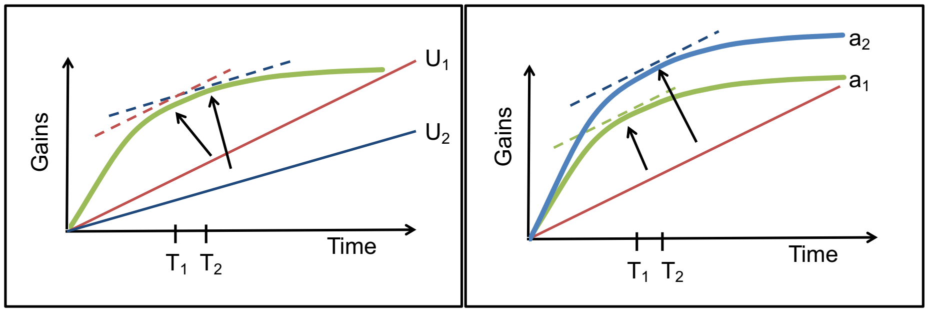

Figure 5. Marginal value theorem allows us to make measurable predictions to task allocation based on both foreground or background changes. (Left) The effect of lowering the lowering the background rate is an increased time in task. (Right) The effect of increasing the foreground rate is an increased time in task.

tangent line to the attainment function where the slope equals $\mathbb{U}^*$. This graphical argument can be used to illustrate what happens to optimal time allocation when the attainment functions differ across tasks or when the average urgency rate changes (see Figure 5.5). These within-task and between-task rate differences result from other variables in the theory, and alone are sufficient to predict time allocation from these otherwise complex relationships.

These graphical arguments emphasize how Equation 5.3 is a relationship between a foreground rate and a background rate, i.e., a comparison between the current task and the environment. We now rigorously extend these relationships.

Quantitative analysis

What we call the average rate of the environment (i.e. background rate) also serves as a *threshold*. This threshold is not intuitive because it is set by an optimization that involves a combination of all the competing tasks, the time between tasks, and the expected yield of the current task. We can separate out these influences by rewriting the main rate equation as a candidate task, and the other tasks lumped into an expected background.

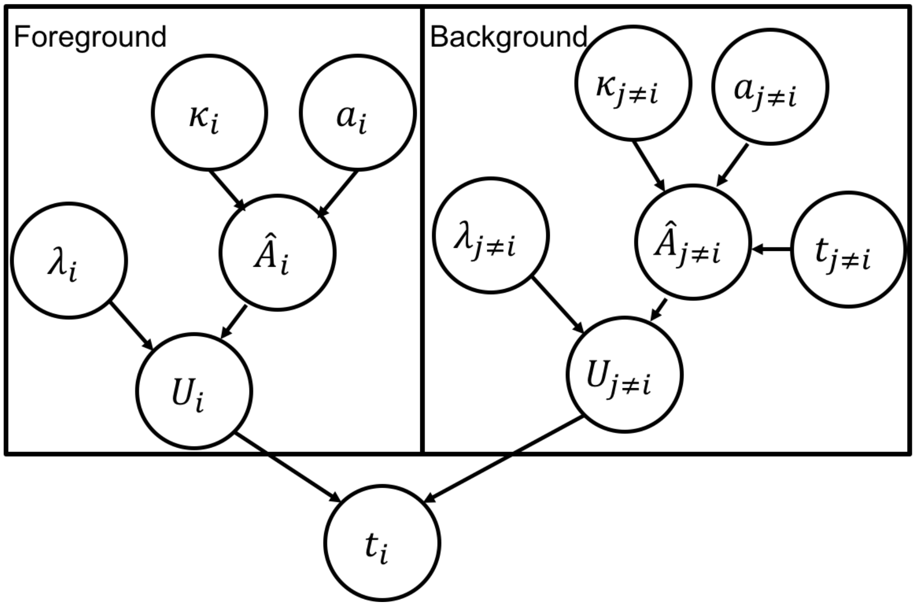

Figure 6. This figure shows the direction of influences on time allocation in the foreground task, separating foreground and background factors. Attainment a_i, availability λ_i, and switch cost κ_i combine to influence the foreground rate U_i for task i. The switch cost of the task and prior attainment combine with whether the task is currently availabile from the current state (task) to determine an expected foreground rate. Similarly, background factors combine to form $U_{i\neq j}$ for all background tasks, which combines with foreground U_i to produce within task time t_i.

$$\mathbb{U}(t_j, t_{-j}) = \frac{\hat{\Lambda}_j \hat{a}_j(t_{w_j})}{\hat{\Lambda}_j t_{w_j} + \sum_{i\neq j} \hat{\Lambda}_i t_{w_i} + T_B} + \frac{\sum_{i\neq j} \hat{\Lambda}_i \hat{a}_i(t_{w_i})}{\hat{\Lambda}_j t_{w_j} + \sum_{i\neq j} \hat{\Lambda}_i t_{w_i} + T_B} \tag{5.4}$$

$$= \frac{a_j(t_{w_j})}{t_{w_j} + L} + \frac{B}{t_{w_j} + L} \tag{5.5}$$

We reduced the above equation by setting $B = (\sum_{i\neq j} \Lambda_i a_i(t_{w_i^*})/\Lambda_j)$ and $L = (\sum_{i\neq j} \Lambda_i t_{w_i^*} + T_B)/\Lambda_j$. This means we can functionally separate out the background influences of our time in a given task, provided we consider the background attainments static (see supplemental 5.9). All factors that affect the magnitude of the desirability of a task have the same, lumped impact on time allocation. Thus we separate these from time-varying components, which have a more complex effect on the time allocation. The overall cost function that we optimize, i.e., the average attainment rate, can be decomposed into three distinct parts:

- Factors that affect the foreground attainment,
- Factors that affect the average background attainment, and
- Factors that affect the rate at which tasks can be engaged in or completed.

Deriving Structural Influences

Here we demonstrate a method to determine constant parametric influences that coincide with the graphical arguments above. We want to know how t^* changes with the parameters $\in [\alpha, \beta, B, L]$. Given the nature of the optimization above, we can use both the implicit function theorem and the envelope theorem to determine the influences of the parameters on the optimal time allocation (see the supplemental 5.9 for details):

$$\frac{\partial t^*_{wj}}{\partial \theta} = -\frac{\partial^2 \mathbb{U}^2}{\partial t^*_{wj} \partial \theta} \tag{5.6}$$

The equation above relates how time in task j changes based on the change in some parameter θ, by relating it to the second derivative of $\mathbb{U}$. We can express this parameter as a function of another parameter, $\theta = z(\rho)$, e.g., if $\theta = B$, then $B = z(\alpha_{-j} \times t_{w-j} + \beta_{-j})$. Then we can use the *chain rule* from the calculus to compute the influence of the nested parameter on the time allocation. For example, if the partial with respect to $\theta = B$ is positive but the partial with respect to $\rho = \alpha_{-j} \times t_{w-j} + \beta_{-j}$ is negative, then the overall product is negative. This gives us a more explicit way of relating the nested parameters in the urgency rate and the time in a task.

Once we look at the influence of all these other components, i.e. B, L, α_j, then the only thing that is left is the complex relationship between the rate of attainment of the foreground task and the time allocated. This depends on the form of the attainment function.

Parameterizing Attainment Functions

Our goal is to derive a set of qualitative relationships between key task parameters like delays, average urgencies, completion rates, etc., and the time spent in foreground and background tasks. To do this we assume that all the attainment functions are quasi-concave with a common parametric form that incorporates these key variables. A simple parametric form for the attainment functions uses a common quasi-concave base function that is scaled, dilated and time-shifted across tasks:

$$a_i(t_{w_i}) = \alpha_i g(s_i t_{w_i} + b_i) \tag{5.7}$$

where $g()$ is the common function that we assume is continuous, and it increases from zero and asymptotes at one; α_i is an overall scale factor that determines the asymptotic attainment; s_i sets the rate of attainment; and b_i sets delays. For example, we could use the sigmoid for $g()$. This parameterized form allows us to incorporate time varying impacts on attainment (such as variable task progress) into $g()$ and static impacts (such as availability λ) into α_i.

Note that the relative urgency between two tasks is determined by comparing the *rates* of the two tasks, equivalent to comparing the time derivative of the two task's $g()$ functions. If the tasks share time delays, then the relative urgency is

$$\begin{aligned} u_i(t)/u_j(t) &= \frac{\alpha_i \frac{dg(s_i t_{w_i}+b_i)}{dt}}{\alpha_j \frac{dg(s_j t_{w_j}+b_j)}{dt}} && (5.8) \\ &\approx \frac{\alpha_i s_i}{\alpha_j s_j} && (5.9) \end{aligned}$$

Thus we can view attainment scaling as directly scaling urgency, or equivalent to urgency scaling. Note that attainment rate has an equivalent effect, and the two can trade off against each other in environments where both vary.

These parameters model the following impacts on time allocation, after plugging the parametric form back into the Urgency equation and computing their differential impacts:

1. *Attainment scaling*: Increasing foreground scale factor α_j increases time in task t_j, while increasing background scale factor α_i decreases it. For illustration, we refer to our paper-writing example, where the author is submitting to a conference. An increase in foreground urgency scaling could come from any factor that incentivizes attending the conference, including the author learning that a distinguished scientist of interest would be attending their talk. This increase would result in the author spending more time on their writing task. On the other hand, a flurry of email notifications combined with the author's hunger, could scale background urgency rate and lower the time spent working on the paper.

2. *Attainment rate scaling (slope)*: For the foreground task, scaling time by $s_i > 1$ increases the slope of the attainment function, while decreasing slope ($0 < s_i < 1$) reduces time in task. Changing the slope of the background has the opposite effect — fast accruing background attainment is tempting and reduces time in task, while a slow background makes the foreground more attractive. In our paper-writing example, the foreground attainment rate, and therefore time-in-task, could be reduced via task conditions such as writer's block or computer difficulties. Background attainment rate increases, e.g., phone calls or text messages, would also reduce the time in the foreground task.

3. *Attainment delay*: Increasing or decreasing the time in task before attainment begins

to accrue relative to average (time shift b_i) makes the task less attractive for greater delays, and more attractive for lower delays. An example of this factor in the paper-writing example is the initial time the author spends figuring out where they left off in the writing process. The more time that is required for the author to start writing again, the less desirable it is to switch into the writing task. Background tasks such as email or television, which require less time to re-engage, become effective distractions.

Impact of deadlines on parameterized attainment function.

While this parameterized form of attainment appears simple, we can relate other possible impacts on the attainment function back to the parameters in equation 5.7. This can be achieved by finding the best-fit equivalent attainment parameters, and then using those equivalent parameters to interpret the impact. Here we explore the impact of deadlines and uncertainty on the parameterized attainment function.

To instantiate deadlines, we multiply the standard attainment function with a Heaviside step function with deadline d, where[4] $h_d(t) = 1_{(-\infty,d)}(t)$. This cuts off any attainment after the deadline time d. Then we use optimization to find an equivalent set of parameters that relate to the deadline, such that $\alpha g(st + b) * h_d(t) \approx \alpha' g(s't + b')$ (for some reasonable error rate). Note we use sigmoids as the attainment's functional form. We plot the resulting attainment functions in Figure 5.7. As shown, the deadlines appear to have the largest impact on the scale parameter α, such that an earlier deadline reduces the overall gain on the equivalent attainment function.

This result appears reasonable, as a sooner deadline should reduce the amount of either instantaneous reward that could be accrued or the likelihood of completing a task. As such,

[4] $1_A(x) = 1$ for $x \in A$ and 0 elsewhere

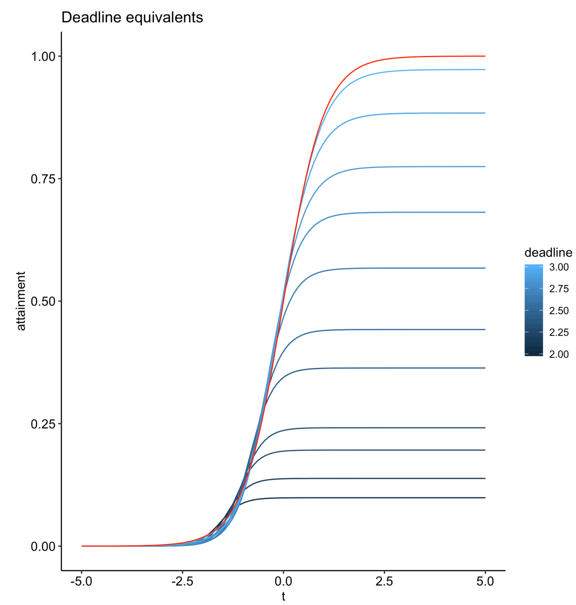

Figure 7. The equivalent impact of deadline on the parameterized form of attainment. The red line (with the largest scale factor) is the baseline attainment function without the deadline. Deadlines function as a step-wise function that decreases attainment to 0 at a certain point.

a sooner deadline functionally reduces the amount of time someone spends on a task in place of alternatives, even possibly causing people to not enter into a task if the deadline looms.

5.5 Psychology of time allocation

Our theory is meant to capture relevant psychological components of time allocation. So far we have derived the mathematical implications of the theory. Here we describe the resulting psychological implications.

5.5.1 Environmental factors that impact time allocation

These quantitative analyses have direct qualitative implications on people's spontaneous task switching based on the structural relationships between parameters. We initially described these in Section 5.3, and summarize our explanations in Table 5.1 and Table 5.2.

For all of the effects on Table 5.1, if the foreground or background urgency rates are markedly different, whichever has the higher urgency will receive all of the time allocation. Thus, changes in background rate have a bifurcating effect - they change the threshold for time allocation, which produces either an obvious choice of foreground over background tasks, *or* a change to time in task. These relationships are complex, but consonant with commonsense reasoning about the conditions that maximize productivity.

An important note is that these environmental factors are due purely to the structural implications of the theory. This contrasts with other psychological perspectives that emphasize the limitations or constraints on human cognition as explanations underlying these phenomena (e.g., Wang, Irwin, Cooper, and Srivastava, 2015). However, an important constraint that people must deal with is in terms of predicting the future given uncertainty. In our theory, attainment is inherently a prediction on the number of relevant evens that are collected. This means it is inherently subject to possible individual differences in ability to infer and predict.

5.5.2 Forecasting and impact of uncertainty

Our theory assumes the agent has some knowledge of the environment's attainment rates. Attainment includes forecasting of relevant events, so an agent must be able to create this forecast. While information from animal learning experiments suggests that they can learn relevant rates of events (Gallistel, 1990), there is always uncertainty in learning these rates.

Table 5.1: Major explanations of phenomena based on time allocation theory, part 1. Continued on Table 5.2. References: **(1)** Draheim, Hicks, and Engle (2016), Mark, Iqbal, Czerwinski, and Johns (2015), Dabbish, Mark, and González (2011), **(2)** Iqbal and Bailey (2008), Graus, Bennett, White, and Horvitz (2016), **(3)** Mark, Czerwinski, and Iqbal (2018), **(4)** Dickey and Meier (2005), Hoffman and Nadelson (2010)

Phenomenon	Domain	Description	Conditions	Explanation	Ref
Distracting environment	Education, IO	Difficulty of engaging in a main task due to distractors	Aversive main task requiring long commitment	Highly available alternatives ($B > a_j$)	1
Interruptions and notifications	HCI	Appearance or reminder of alternative task can cause immediate switching	Main task needs continuous work	Sudden change in background rate ($\Uparrow \lambda_i \rightarrow \Uparrow B$)	2
Dual effect of task blocking	HCI	Low completers benefit from difficult transitions, while high completers do not	Total attainment for high completers is more impacted by blocking	Blocking reduces background λ_i, and $\Downarrow \lambda_i \rightarrow \Downarrow B$. BUT $\Downarrow B \rightarrow \Downarrow \mathbb{U}$, which can reduce overall time in tasks (if attainment is low).	3
Gamification and accountability	HCI, Social	Some gamifications, and added social accountability, can improve time on task	Points or accountability closely track task progress	Access to task progress signals increase attainment	4

Table 5.2: Major explanations of phenomena based on time allocation theory, part 2. Continued from Table 5.1. References: **(5)** Carver (2003), **(6)** Koo and Fishbach (2012), Schmidt and Dolis (2009), **(7)** Hartonen and Alava (2013), Jarmolowicz, Hayashi, and Pipkin (2010)

Phenomenon	**Domain**	**Description**	**Conditions**	**Explanation**	**Ref**
Quitting near task completion	Social	Task requires completion to be satisfied, but people quit early.	Task has diminishing returns.	Time to complete is less than optimal time	5
Progress sweet spot (dual effect of progress)	Social	Both high and low progress can cause switching	Presence of alternative tasks	Declining rate of attainment (progress) can cause switching, while an overall low attainment will produce little time in task	6
Dual effects of deadlines	Social, IO	Deadlines cause switching when far but sticking when near	Farther deadlines have harder to estimate impact on completion	Changing uncertainty of finishing: higher κ as deadline approaches.	7

This means attainment incorporates whether events can be achieved, consonant with goal feasibility (Gollwitzer, 1990).

Uncertainty impacts the attainment function in interesting ways. We can show this impact by relating uncertainty in our parameters from equation 5.7, α, s, and b, to the equivalent impact on these same parameters *without* uncertainty. In other words, if we increase uncertainty in the scale of the attainment function, what equivalent impact is that on the parameterization? From this, we can then deduce the resulting impact on time allocation.

To explore the impact of uncertainty, we allow each parameter to be a random variable before taking an expectation. Then we find the best-fit equivalent parameterization of the same function to find the equivalent impact on the parameter values that an increase in uncertainty produces. For example, to find the impact of uncertainty on slope, we allow $s \sim Gamma(\mu, \sigma^2)^5$, and then take expectation over s, and find a closed fit set of parameters α', s', b' such that $E_s[\alpha g(st + b)] \approx \alpha' g(s't + b')$. We can then vary the uncertainty in s, and observe the impact on the various parameters.

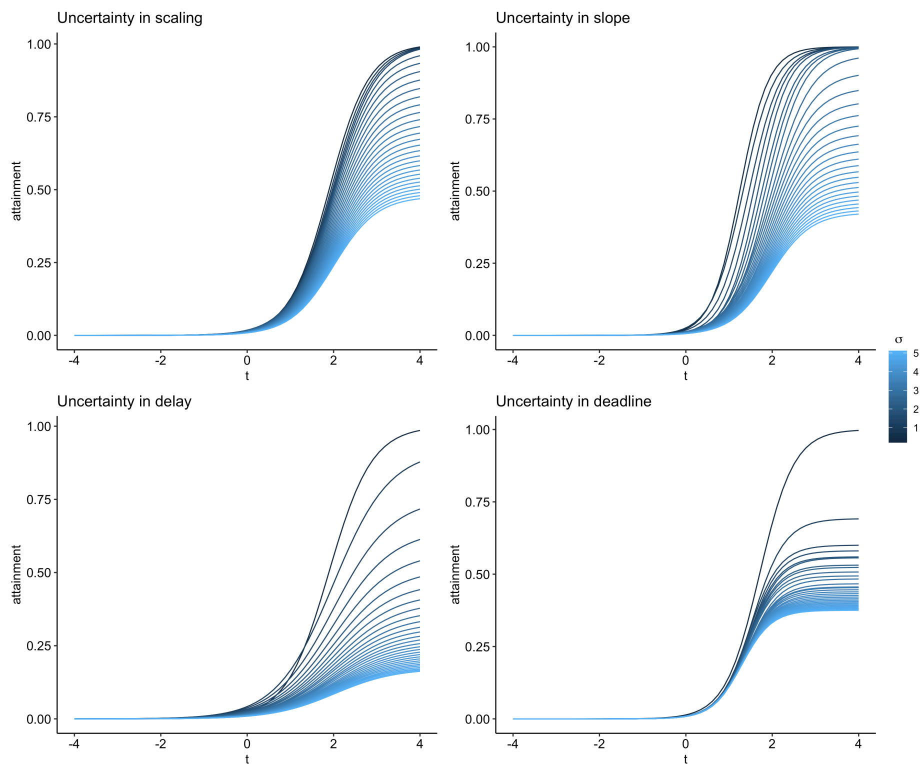

Figure 5.8: The functional impact of uncertainty on a given parameter on the overall attainment shape. Here we have the scale α, slope s, delay b and deadline d impacted by uncertainty in the parameter. α, s and b are distributed via gamma, and deadline is distributed normally. The other parameters are kept otherwise constant arbitrary values (2, with mean values at 2) for comparison.

On Figure 5.8 we plot these best-fit attainment functions as we increase uncertainty. While uncertainty can impact all parameters, the result of uncertainty appears to have the

[5] Here the gamma distribution is re-parametrized by it's mean and standard deviation, such that shape $k = \mu^2/\sigma^2$ and scale $\theta = \sigma^2/\mu$

strongest impact on the overall attainment scale. Increase in uncertainty functionally decreases the scale of the attainment (see Figure 5.9).

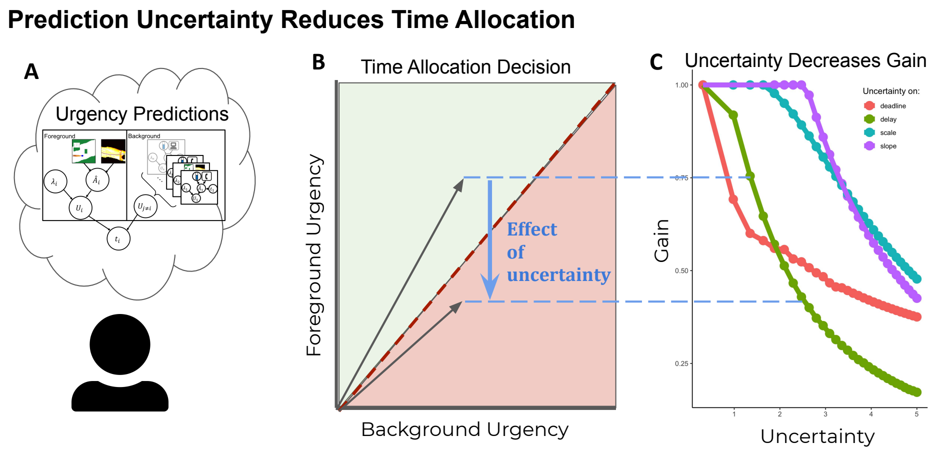

Figure 5.9: **A)** Time allocation decisions involve predictive forecasts of the key structural parameters, including deadlines, availability and completion rates. These predictions coalesce into overall urgency forecasts for candidate foreground and the foreground-relative background tasks, which compete to determine time allocation. **B)** Time allocation decisions can be visualized as competition between foreground and background urgency. Foreground tasks with urgency greater than background (green shaded region) are rationally best to allocate time to, while tasks with lower than background urgency should not be engaged in. Across all major structural factors, the effect of uncertainty is to lower urgency (blue arrow). **C)** The relationship between uncertainty (increase in standard deviation) and the gain on the attainment functions (scale parameter α). In all cases, gain decreases at increasing uncertainty with different curves that are likely based on the impact the parameter has on the overall shape of the function. The effect of uncertainty results from the linear relationship between the cumulative and rate attainment functions; a constant decrease in attainment (due to constant decrease in gain) results in a constant decrease in rate.

The impact of uncertainty can be seen clearly in the marshmallow delayed gratification task (Kidd, Palmeri, & Aslin, 2013). We can predict when a child will decide to stick with the current task (wait for two marshmallows) or switch to an alternative (consume the one). The time discounting in this task functions as an attainment delay as in equation 5.7. When children are placed in an unreliable situation, it functionally increases their uncertainty as to the delay time. As shown in Figure 5.9, an increase in the uncertainty of the delay will cause an overall decrease in an attainment (shown as the effect of uncertainty). This produces a reduction in the gains as shown in Figure 5.5. This delays the optimal switching point, so that $t_{w2} < t_{w1}$ and the child will quit earlier than otherwise in the waiting task.

Humans almost certainly estimate the quality of the environment (Eldar, Rutledge, Dolan, & Niv, 2016) and likely also learn the value of different tasks and goals by experience (Srivastava & Schrater, 2015). This can result in different prior expectations on the attainment function. While we have emphasized how otherwise seemingly inappropriate switching can occur due to rational time allocation, cognitive constraints (e.g., computational costs or limitations as in Lieder and Griffiths, 2019) such as in appropriately forecasting attainment, can cause inappropriate switching. Computational constraints in inference likely play a role in how people generally infer attainment, even if they are approximately correct, such as the use of the availability heuristic (Tversky & Kahneman, 1973).

Moreover, someone with a poorer ability to predict any aspect of the attainment function will result in them switching preemptively *as if* the task has a lower overall value. Someone with a difficulty of predicting the attainment on most tasks will appear to gravitate to tasks with a more certain attainment function, simply due to spending more time in tasks with less uncertainty, and uncertainty possibly dropping a task below the background urgency level before task engagement. Differences in expectations or the uncertainty on these parameters are relevant targets for computational psychiatry. The relationships between time allocation and autism spectrum disorder (Sinha et al., 2014) or attention deficit hyperactive disorder (Hauser, Fiore, Moutoussis, & Dolan, 2016) can be possibly be due to a different prior on, or high uncertainty in, the attainment function.

5.5.3 Which events make up attainment?

Our scheduling theory is reliant on the attainment function, $a_i(t_{wi})$. While we have described the mathematical constraints such a function requires (e.g., saturation), we have not fully described what real-world events it might correspond to. Importantly, it is a forecast over any events of interest. Within foraging theory, these events are generally food consumption with associated energetic gains (Stephens & Krebs, 1986). We have allowed our attainment function to represent more general events of interest for psychologists.

Attainment can include qualitatively different kinds of events, including internal and external events. An important interpretation of equation 5.7 is that it represents what's called *linear scalarization*. Scalarization is a method in multi-objective optimization where multiple (independent) objective functions $f_i(x)$ are combined to produce a single objective $F(x)$. In the linear form, they are weighted together: $F(x) = \sum w_i f_i(x)$. In equation 5.7, the scale factor α_i functions as a weighting across possibly distinct task attainment functions $g_i()$. This allows us to incorporate both internal and external events across attainment functions. The relative difference in weights can be significant in understanding individual differences in time use.

Externally rewarding events provide the most obvious source of attainment. While we have pushed against a solely reward-based function, there are some clear external events that provide task engagement due to rewards. The most obvious example is food consumption in humans, along with many other rewards driven by basic biophysical systems (Hayden, Pearson, & Platt, 2011; Berridge, 2004). Casino-style games also provide clear, external events with presumed value (Schüll & Library., 2012). However, extrinsic rewards are fundamentally limiting in explaining human task engagement (Ryan & Deci, 2000).

As previously mentioned, task progress itself can be relevant to human task engagement

(Carver, 2003). Much of human behavior is directed towards constructed goals with little explicit reward attached even for finishing. Goal progress per se can be an attainable event that people track, that is, signals indicating how close one is to achieving a goal. These events can be both external and internal. For example, while word count can be structured as a motive, people also have a more nebulous sense of progress in writing. This can be formulated as change in uncertainty of goal satisfaction, or a change in entropy.

Intrinsic motives, such as curiosity, mastery, and exploration, refer to the desire to engage in a task not due to external events that occur after task completion (such as rewards), but due to the nature of the task itself. Video games, sports, and art all represent examples of tasks engaged in for these intrinsic reasons (Ryan & Deci, 2000). Importantly, the type of internal events captured can qualitatively change the motives involved. Human behavior is often intrinsically motivated by a desire to learn (Loewenstein, 1994; Litman, 2005; Schmidhuber, 2010), master and control the environment (Ryan & Deci, 2000; Csikszentmihalyi, 1990), and explore (Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Cohen, McClure, & Yu, 2007). Many of these refer to internal states of an agent, or the relationship between the agents perceptions, actions, and the world. Mastery, for example, is a reduction in uncertainty of output competence (i.e., actions more likely produce the desired effects); this has been formalized as an increase in empowerment (Klyubin et al., 2005). Incorporating these types of measures into attainment can help clarify how different types of events are weighted.

Intrinsic motivation measures are typically defined in information-theoretic terms (Oudeyer & Kaplan, 2007, November). For example, the entropy rate, $H(x_t)$, provides a measure of how much information is carried by a time-varying signal x_t. If x_t is a relevant state for an agent, for instance, knowledge states of the world, then a reduction in uncertainty can carry intrinsic value. If we consider events where the reduction in entropy $\Delta H =$

$[H(x_{t-1}) - H(x_t)]$ is greater than some threshold: $\Delta H > \epsilon$, these events can themselves provide task engagement via a measure which captures the incentive value of information.

Empowerment as an information-theoretic quantity measures an agent's capacity to predict its future given its current actions. Empowerment measures the amount of information an agent's actions can put into the environment (as measured by the states achievable). In other words, given a particular state, empowerment is a measure of the both the number of future states available as well as the agency of the agent's actions for achieving those states. Given the agent's state is x_t at time-step t, the agent's actions over the next n time-steps is $A_t^n = \{a_t, \dots, a_{t+n}\}$ and the future states are $x_t^n = \{x_t, \dots, x_{t+n}\}$, the rate of information I that can be transmitted along a channel between actions and future states is the mutual information between them: $I(A_t^n; X_t^n)$. Empowerment $\mathfrak{E}_t(x_t)$ is the highest information rate achievable over a horizon n from the current state:

$$\mathfrak{E}_t(x_t) = \max_{p(A_t^n)} I(A_t^n; X_t^n) \tag{5.10}$$

Integrating empowerment into time allocation

For a video game, a player's empowerment rate can be directly used to predict their engagement with a task by integrating empowerment back into our original theory. We can insert empowerment from equation 5.10 into equation 5.7, such that $a_i(t_i) = \mathfrak{E}_t$, i.e.:

$$a_i(t_{wi}) = \alpha_i \int_0^t \mathfrak{E}_t \tag{5.11}$$

where α_i acts as a scaling factor that weights the attainment for satisfying the goal state (which is specified by the weighting factor w). This new $a_i(t_{wi})$ functions as explained within the above analysis — we construct the urgency $\mathbb{U}$ from taking a weighted average rate across tasks (equation 5.2) which allows us to use the mean value theorem (equation 5.3).

We can also derive a special case when empowerment is the dominant form of attainment for all available tasks. This can be the situation when multiple video games are traded off with one another, in which case the empowerment measure can be solely used to predict time allocation. In that case, equation 5.3 results in:

$$\alpha_i \mathfrak{E}_{t,j} = \mathbb{U}^* \tag{5.12}$$

Since equation 5.11 is the cumulative empowerment, taking the derivative results in the interesting conclusion that the current empowerment (scaled and weighted) is what determines quitting (see Figure 5.10). When the current empowerment drops below the average urgency, quit.

Using empowerment as the primary attainment function for a task allows us to predict the engagement associated with videogames, an important family of tasks known with the potential to dominate time-allocation. We show the empowerment at each state of a helicopter game (based on an expert player) as a heatmap — accumulating empowerment behaves exactly like accumulating an external reward rate, with important psychological consequences. Our theory predicts that uncertainty in the reward rate will generically decrease time allocation. For task domains where external reward events are sparse, we predict that scheduling will be dominated by intrinsic motive event rates like empowerment, which can be dense, especially in familiar games where empowerment rates can be accurately learned.

Psychologically, allowing both intrinsic and extrinsic event rates to compete provides interesting rational explanations for the addictive potential of video games. Consider the simple example of our helicopter game mentioned in Figure 5.10, where players learn to fly a small helicopter through a procedural tunnel, navigating around obstacles. Drops in computed empowerment often correspond directly to common situations where people quit,

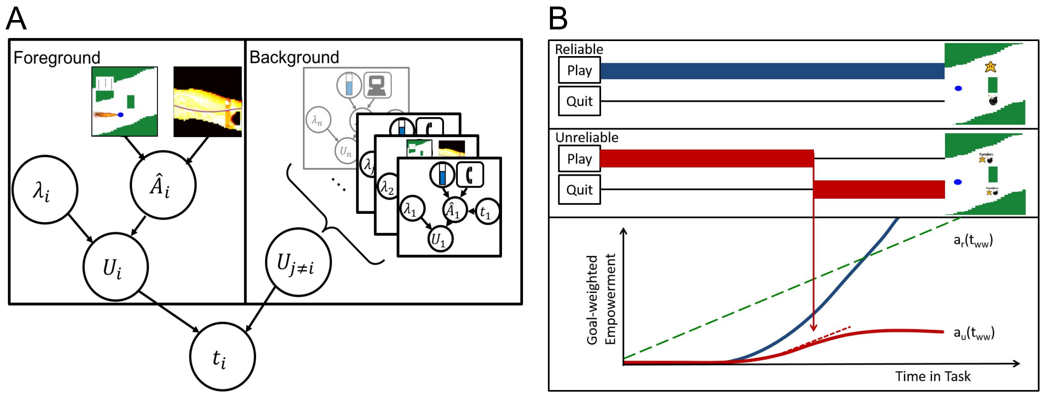

Figure 5.10: Using empowerment as our attainment quantity to predict time in a video game. **(A)** We revisit Figure 5.6 and show how it applies to playing a video game motivated by mastery (empowerment). One approach is to treat a simple helicopter flying game as the foreground task, and other activities as competing background tasks. We show computed empowerment in this game, using the approach in (Mohamed & Rezende, 2015) to compute attainment in each state (colored). These combine with other factors to produce foreground urgency. Each available background task $1, ..., n$ (shown as plates) average together to produce the overall background attainment rate $U_{j\neq i}$. **(B)** The time in task for a reliable game environment vs the time in task for an unreliable game environment, along with the corresponding attainment curves based on goal-weighted empowerment (Edge, 2013). Notice that the lower rate of empowerment for the unreliable task induces a shorter time in task, while it takes longer for the empowerment rate in the reliable task to drop to the environment average.

e.g., right before dying (when outcomes are certain). It is also straightforward to imagine how changing the game will impact time played via empowerment. Having an extensive laborious menu to restart or quit will keep people playing, but after leaving, players might not return as often.

Certain features in games that currently exist can also be understood given the integration of empowerment and time allocation theory. Many games, such as League of Legends ("/Remake FAQ for League of Legends", 2016), will penalize people for quitting a game early; if a person recognizes a drop in empowerment due to a known outcome (e.g., a long, drawn-out loss), they'll quit. These punishments for quitting can be needed to incentivize fair play. Interestingly, the most common penalty is a long wait-time before playing the game again. By contrast, players often have the option to restart the game early on, again if the outcomes of the game seem certain or due to poor luck. Restarting can help reset empowerment. However, all these effects are managed based on how empowerment rate compares with background.

5.6 Empirical measures of attainment

So far we have discussed attainment in terms of the psychological modeling and interpretations, and have shown that many of our predictions do not rely on a particular measure for the attainment function. This leaves the question of how do we pragmatically measure attainment to make time allocation predictions. We have previously shown how we can incorporate a priori attainment measures, as in empowerement. Here, we discuss methods of empirically measuring attainment.

Our $\hat{a}$ functions can be elicited from time allocation and choice behavior, similar to eliciting utility functions in behavioral economics and decision psychology (Hastie & Dawes, 2010). Eliciting $\hat{a}$ from the allocated time provides a useful description of the preference for a task. However, to obtain a dynamic attainment function, the indifference point for a task and all alternatives (e.g., monitory reward) would have to be measured for each time point in the task, making this process rather infeasible for general use, but possible for tasks with extrinsic rewards upon completion.

One possibility is to estimate a parametric form of attainment using self-report, such as via sampling methods or through time diary approaches (i.e., retrospective surveys and interviews such as United States., 2003). Research on psychological flow (Csikszentmihalyi & Lefevre, 1989), task-unrelated-thoughts (TUTs) (Kane, Kwapil, Mcvay, & Myin-germeys, 2007), and activity analysis (Aggarwal & Ryoo, 2011) use experience sampling methods to measure preferences for various real-world tasks. The difficulty with this approach, however, is the relationship between the subjectively experienced value of a task, the retrospective or remembered appraisal, and the future predictions reported by subjects can be incongruous (Kahneman & Krueger, 2006b). Rather than using subjective appraisal, the $\hat{a}$ function could be parametrically estimated by observing the pattern of a user's switching behavior.

Behaviorally, high $\hat{a}$ tasks should produce engagement. Predicting when an individual disengages from a task can be done by using survival analysis (Aalen, Borgan, & Gjessing, 2008), a statistical method used to analyze the time or duration till a given event occurs, such as patient death or machine failure. Previous work used survival analysis to predict when an individual gets bored of and quits listening to a song (Kapoor, Subbian, Srivastava, & Schrater, 2015), which can be easily extended to predict general task quitting. Survival analysis often involves specifying a *hazard function*, which is the instantaneous rate of an event occurring. While the hazard function is often estimated just by using event times, regression methods such as the Cox hazard model can be used to determine what factors best predict those times, such as quitting a video game. Researchers have used survival analysis and Bayesian optimization to adaptively change game parameters (Isaksen, Gopstein, & Nealen, 2015; Khajah, Roads, Lindsey, Liu, & Mozer, 2016). However, predicting what parameters in a game are important for engagement or difficulty (and should be manipulated) is challenging, and generally requires an experienced designer to intuit. In general it is more satisfying to have diverse measures of $\hat{a}$ than solely switching behavior.

The $\hat{a}$ function is an integrated urgency signal that should produce measurable correlates in behavior, psychophysiological responses and neural activity. Because allocating time and resources to a task should have neural and psychophysiological consequences, we believe that biophysical measures of attention and arousal may be useful to monitor engagement, focus, boredom and disengagement (e.g. D'Mello and Graesser (2012)). This could be done by operationalizing attainment directly through associating it with measures of attention or arousal. In particular, this could provide a measure of urgency for the task without interrupting the subject (which subjective reports require), or modeling the current task. However this approach requires more work detailing the precise relationships between various measures and time in task.

An integrative framework is to predict $P(t_{quit}|\hat{a}_i, \hat{a}_{back})$ by using a parametric survival function, such as a Gamma or Wald distribution, with a mean value that is a function of both the $\hat{a}_i$ and $\hat{a}_{back}$, as related by our optimal time allocation theory in Figure 5.6. Different covariates such as heart rate can be regressed by parametrically modeling $\hat{a}_i$ using equation 5.7. Then the current time allocation theory allows us to predict the *direction of influence* of different factors which, when combined with survival analysis, allows us to precisely predict time allocation. This could then be used to perform credit assignment, to determine which signals correspond to attainment and be subsequently used elsewhere to predict time on task.

5.7 Comparison to alternative frameworks for time allocation

Here we discuss how our framework relates to alternative perspectives for time allocation, in particular self-regulation and reinforcement learning.

5.7.1 Self-regulation and Control Theory

Models of self-regulation are often described using the language of control theory (Carver & Scheier, 1981, 1998). The original model of self-regulation from Carver and Scheier (1998) takes the form of a T.O.T.E. (test, operate, test, exit) feedback control system from Miller, Galanter, and Pribram (1960), a classic cybernetic architecture. These regulation control theories define a goal as a set point to be either achieved or avoided, a regulation process whereby the currently perceived state is compared with the desired goal, and actions which are taken to reduce any discrepancy (Vancouver, Weinhardt, & Schmidt, 2010). While time allocation can result from these control process, self regulation has not been framed as a scheduling problem. In scheduling problems, time is the critical decision variable and is the direct locus of control. By contrast, self-regulation theory has difficulty accounting for time allocation between goals, as it focuses on *progress toward a goal*.

Generally speaking, self-regulation is concerned with the regulation of a particular goal rather than dealing with multiple goals; however, recent research has recognized that these theories need to address multi-goal pursuit (Neal, Ballard, & Vancouver, 2017), to account for trade-offs in time allocation between competing tasks. For instance, in Ballard, Vancouver, and Neal (2018) they model multiple goal pursuit as a more traditional decision problem, where an agent selects between pursuing goals on the basis of their expected utility. To explain switching behavior, they construct an ad hoc utility function that integrates goal completion, importance, and time pressure into an expected utility term which can be updated during performance of a task, producing a reactive agent with greedy goal selection. This approach is a standard way to reduce sequential dynamic decision problems into simple, static but myopic decisions — and it fails to account for look-ahead and planning processes essential for more sophisticated behavior. In contrast, we compute the future expected value obtained by a look-ahead, and use that to derive both the form of the utility functions and the structural properties of the trade-off. Our approach represents a rational multi-goal agent with the ability to forecast and plan, and subsumes the simpler approach as a special case with a short planning horizon.

5.7.2 Reinforcement learning and hierarchy in cognitive control

Reinforcement learning (RL) is a family of methods for learning policies, or maps between states and actions, that maximize reward with respect to a goal (Sutton & Barto, 1998). RL has enjoyed substantial success in modeling different components of the brain's reward and decision making systems (e.g., McClure, Berns, and Montague, 2003; Daw, Niv, and Dayan, 2005; Tanaka et al., 2004). However, standard RL is directed at selected actions towards a single goal (instantiated via a reward function). There is abundant evidence for hierarchy both anatomically (Badre, 2008; Botvinick, 2008; Koechlin, Ody, & Kouneiher, 2003) and computationally (Todorov, Li, & Pan, 2005; Collins & Koechlin, 2012). We believe this hierarchy has an interpretation as a goal-level controller, as emphasized by self-regulation theory. The computational problem of multiple goals has led to increasing focus on hierarchical forms of RL (Sutton, Precup, & Singh, 1999; Botvinick, Niv, & Barto, 2009).

Hierarchical RL (hRL) treats multi-goal pursuit through the lens of switching between policies, where individual policies can be tailored to be solutions for distinct goals (Sukhbaatar, Denton, Szlam, & Fergus, 2018). Higher order control then learns a hierarchy of subpolicies, treating subpolicy selection as a higher order action (Bacon, Harb, & Precup, 2017). In principle, hierarchical RL systems can model task switching, and we can relate our approach to a family of hRL problems involving a set of mutually exclusive goals, partially observable event observations, intrinsic rewards, transit costs and an average reward rate criterion. At the time of writing, this conjunction requires a non-trivial extension of what's currently achievable in hRL (e.g., Sukhbaatar, Denton, Szlam, and Fergus, 2018; Rafati and Noelle, 2019; Kulkarni, Narasimhan, Saeedi, and Tenenbaum, 2016; Vezhnevets et al., 2017; Bacon, Harb, and Precup, 2017; Nachum, Gu, Lee, and Levine, 2018). More importantly, even if we were to solve the problem using hRL, the structural properties of the solution would be opaquely embedded in the system.

Our approach is to take advantage of the simplifications afforded by a few simple assumptions about tasks, their dynamics, and their rewarding properties to derive detailed structural predictions. These predictions have more scientific value as they pinpoint essential relationships while abstracting away the particular domain features that would be embedded in any hRL modeling approach. The value of hRL will be in modeling particular domains, where the high-level assumptions may be violated, and resulting behavior is difficult to predict. It could also help to expand our approach to include online learning, given future advances in model-based hRL approaches.

5.8 Discussion

Summary of Contributions

In this paper, we reformulate task switching as the control problem of time allocation. Our approach allows for a more ecologically valid understanding of human behavior by focusing on the optimal stopping problem that humans are solving during task selection.

We introduce the concept of urgency, whose key structural factors from the task and environment influence switching. We then detail how to integrate an example attainment signal: empowerment. Empowerment can explain the importance of a person's intrinsic motivational factors when determining their attraction to games and other activities that are void of real-world importance and often directly replace tasks of measurable extrinsic value. We also explain how these urgencies can possibly be elicited from raw task switching data.

We enrich the standard foraging framework by augmenting the foraging agent giving them 1) beliefs about task and goal availability 2) beliefs about task completion due to activity 3) intrinsic urgency signal, and 4) dependencies between tasks. We have shown that despite this added complexity, we arrive at a similar result to optimal foraging theory. The marginal value theorem from this theory provides a simple condition that time allocation must also obey. Our work provides an alternative interpretation to the results of many of the standard task switching phenomena as in Table 5.1; human task switching is a natural implication of solving optimal time allocation.

This key idea impacts many domains: essentially anywhere that individuals must allocate time between different tasks. Some examples include persistence of employees and students at writing tasks (Rosen, Mark Carrier, & Cheever, 2013), individuals on exercise and dieting regimens (Mata et al., 2009), impulsive spending habits (Baumeister, 2002), and the distractability of phone or computer applications (Marulanda-Carter & Jackson, 2012; Gazzaley & Rosen, 2016). Media multitasking provides a prototypical task switching situation, and despite the recent broad concern on multitasking's effects on cognition and attentional processes Van Der Schuur, Baumgartner, Sumter, and Valkenburg (2015), the underlying theories for when and why people switch is generally lacking (Wang, Irwin, Cooper, & Srivastava, 2015). Our results here provide an overall framework to consider how the foreground task and background environmental factors integrate to produce situations conducive to high multitasking.

Interpreting switching behavior: cost of switching

Importantly, cognitive switching costs play a critical role in alternative theories of time use Wang, Irwin, Cooper, and Srivastava (2015). However, while we incorporate switching costs formally into our theory, we do not need switch costs for any of the major phenomena we explain. Thinking of these costs as a distinctive and critical component to explain real-world spontaneous task switching may be a mistake. While they can certainly be there

(Rosen, Mark Carrier, & Cheever, 2013) and can be important to explain some of the negative consequences of switching, they are not needed per se. Instead, the core phenomena can be explained based upon the structural aspects of time allocation.

Interpreting switching behavior: Individual differences and the role of uncertainty

We have previously discussed that people must monitor attainment, but how do they? It is easier to have reliable signals when there are frequent events. When the only events that people can monitor are task completion events, as in many modern work tasks, people will not get a good estimate of their rate of attainment. In these cases when environmental feedback rates are infrequent, intrinsic feedback rates are likely to dominate. Intrinsic rates, such as those driven by empowerment or other internal events, can drive behavior since they are always available. Since uncertainty leads to a reduction in attainment, if there are multiple components of attainment the reliable ones should win. This means the set of tasks that are available can strongly impact time use, due to the trade-offs from background urgency.

Interpreting switching behavior: Role of the environment

The impact of background urgency also speaks to the importance of environment factors on task engagement. Some experiments (Alexander, Coambs, & Hadaway, 1978; Xu, Hou, Gao, He, & Zhang, 2007) provide evidence on the importance of alternative tasks in evaluating addiction. While addiction has a clear neurochemical basis, it is possible that environmental "background rate" factors can modulate its frequency. The design of gambling casinos appears to take explicit advantage of manipulating the perceived availability of alternative tasks in an attempt to maximize time-on-task (Schüll & Library., 2012). A similar view shows up in treating app-addiction as a defect in cognitive control, when in reality it could easily be functioning as optimal time allocation. Application designers have a big incentive to tap the natural motivational systems of their users to maximize their app's urgency, but as the urgency of a population of applications increases, the increased background urgency rate results in *a decrease in the average time in task*. This is general, in that any high background rate environment will produce high switching.

The environment also provides a set of time constraints and deadlines which must be considered when allocating our time. While our approach is focused on spontaneous switching in time allocation, we are able to incorporate the impact that time constraints should have on time allocation. A more complete analysis of the effect of deadlines is definitely needed — much of time allocation is explicitly organized, especially in work or education domains: e.g., waking up with an alarm and going to a bus stop to get to work on time. Time constraints not only shape our behavior, they also place distinctive memory demands on a cognitive architecture which needs to plan for future goal availability in light of the time constraints. During planning, both future goals and time constraints need to be integrated into prospective memory (Einstein & McDaniel, 2005), for example, by filtering out prospective goals which are infeasible. Constraint-sensitive remembering of prospective goals can be thought of as generating an internal cue to a shifting background rate of availability.

However, the general use of clocks or calendars suggests that any such filtering process is quite error-prone – we very often preferentially make meta-level decisions to allow external cues to regulate our time allocation– an "opt-in" pre-commitment strategy. It is also clear that many situations challenge pre-commitment approaches, creating difficult clock-based switches between behaviors, especially if the urgency of the task switched away from is high. While it's tempting to view this as an alternative decision process that competes with an internal implicit scheduling process, we've shown that deadlines and time constraints can be incorporated into the predictions of our implicit scheduling system. These explicit decisions could be thought of as attempts to manipulate our implicit scheduling system through changing information or cues, rather than overriding the decision process. These possibilities should make distinct predictions which can be decided by future research, as a reliability comparison should afford different interventions than an overriding process (i.e., cognitive control).

5.9 Conclusion

We introduce a rational theory of spontaneous task switching that makes broad qualitative predictions consonant with major findings. Our approach uses a computational rational analysis, predicting human behavior by specifying the problem we believe they are solving (Anderson, 1990). As such, our theory does not directly specify the algorithmic solutions or neural implementation of the problem. This means we are predicting human behavior but not necessarily how we believe humans are producing the solution. In order to link this theory with a mechanistic and neural understanding of human engagement, we can begin to relate the structural theory to neural circuits which could implement the marginal value rule and priority cues. For example, previous work has indicated that animals could use a drift-diffusion style computation to estimate within-task reward (Hayden, Pearson, & Platt, 2011) (where reward is related to attainment). Our theory makes novel predictions for the interactions between task representations and for additional circuitry needed to maintain and update metacognitive task priority. Future work in this area would forge important links with other research in decision neuroscience and computational psychiatry.

By reinterpreting human spontaneous task switching in terms of a more ecologically rational approach, we achieve a better, more holistic understanding of human task switching behavior. It also suggests new paths to change that behavior through environmental changes and uncertainty reduction, depathologizing individual differences in task switching. If we understand the natural problems that humans face and solve, we can extend laboratory-based research to more practical real-world applications.

Supplemental

Optimal time allocation derivation

Symbol	Definition
$\mathbb{U}$	Average net urgency rate
A	Average net attainment gain
T_B	Average time between tasks
T_w	Average time within tasks
$a_i(\cdot)$	Net gain in task i as a function of the time within task i
t_{wi}	Time within task i
λ_{ij}	Availability of task j from task i
$\kappa_i(\cdot)$	Switching cost for task i as a function of the time in task i and j

Here we work out the optimal time allocation result. We first define a set of notation in table 5.9. Initially, we start out with the standard equation for average gain rate from equation 5.1.

$$A = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_{ij} T_B a_j(t_{wj}) \kappa_j(t_{wj}|t_{wi})$$

where λ_{ij} is the conditional probability of entering j from i (so $\lambda_{ij}T_B$ is the rate). Both $a_j(t_{wj})$ and $\kappa_j(t_{wj}|t_{wi})$ multiply to determine the overall net value of task j. To produce an average gain rate, assuming we are currently in task i, we then just sum up:

$$A_i = \sum_{j=1}^{n} \lambda_{ij} T_B a_j(t_{wj}) \kappa_j(t_{wj}|t_{wi})$$

and we produce the total average gain rate by summing across i, so $A = \sum_{i=1}^{n} A_i$. This means that the total average reward rate is:

$$\mathbb{U}(t_{w1}, t_{w2}, ..., t_{wn}) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_{ij} T_B a_j(t_{wj}) \kappa_j(t_{wj}|t_{wi})}{\sum_{i=1}^{n} \lambda_{ij} t_{wj} + T_B} \tag{5.13}$$

However, in order to optimize, we need to take a derivative with respect to t_{wj}, such that it does not depend on t_{wi}. There are different ways of doing this. One is to remove the notion of lost progress, or in a sense, remove the agent's prediction of lost progress. So then $\kappa_j(t_{wj}|t_{wi}) = \kappa_j(t_{wj})$. Then one can bring the summation over i in and bring T_B out:

$$\begin{aligned} A &= \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_{ij} T_B a_j(t_{wj}) \kappa_j(t_{wj}) \\ &= \sum_{j=1}^{n} (\sum_{i=1}^{n} \lambda_{ij} T_B) a_j(t_{wj}) \kappa_j(t_{wj}) \\ &= \sum_{j=1}^{n} \hat{\Lambda}_j a_j(t_{wj}) \kappa_j(t_{wj}) \\ &= \sum_{j=1}^{n} \hat{\Lambda}_j \hat{a}_j(t_{wj}) \end{aligned}$$

where $\hat{\Lambda}_j$ is the average availability of task j across tasks i, and $\hat{a}_j(t_{wj}) = a_j(t_{wj})\kappa_j(t_{wj})$. Given that both a_j and κ_j are monotonically increasing functions, this means that $\hat{a}_j$ is as well.

Another similar method of removing the dependence on t_{wi} is to instead consider the worst-case time spent away, t_{wi}^*, and keep it constant, thereby also reducing $\kappa_j(t_{wj}|t_{wi}) = \kappa_j(t_{wj})$, and allowing for the above substitution.

Now, returning to equation 5.13, we can replace A with our simplified A equation:

$$\begin{aligned}\mathbb{U}(t_{w1}, t_{w2}, ..., t_{wn}) &= \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_{ij}T_B(t_{wj})\kappa_j(t_{wj}|t_{wi})}{\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_{ij}t_{wj} + T_B} \\ &= \frac{\sum_{j=1}^{n}\hat{\Lambda}_j\hat{a}_j(t_{wj})}{\sum_{i=j}^{n}\sum_{i=1}^{n}\lambda_{ij}t_{wj} + T_B} \\ &= \frac{\sum_{j=1}^{n}\hat{\Lambda}_j\hat{a}_j(t_{wj})}{\sum_{i=j}^{n}\hat{\Lambda}_j t_{wj} + T_B}\end{aligned}$$

In order to take the derivative with respect to an arbitrary t_{wj}, we separate out all other variables in the equation:

$$\mathbb{U} = \frac{\hat{\Lambda}_j\hat{a}_j(t_{wj}) + \sum_{i\neq j}\hat{\Lambda}_i\hat{a}_i(t_{wi})}{\hat{\Lambda}_j t_{wj} + \sum_{i\neq j}\hat{\Lambda}_i t_{wi} + T_B} = \frac{\hat{\Lambda}_j\hat{a}_j(t_{wj}) + a_j}{\hat{\Lambda}_j t_{wj} + b_j}$$

where $a_j = \sum_{i\neq j}\hat{\Lambda}_i\hat{a}_i(t_{wi})$ and $b_j = \sum_{i\neq j}\hat{\Lambda}_i t_{wi} + 1$. Taking the derivative, we get:

$$\frac{\partial R}{\partial t_{wj}} = \frac{\hat{\Lambda}_j\hat{a}'_j(t_{wj})(\hat{\Lambda}_j t_{wj} + b_j) - (\hat{\Lambda}_j\hat{a}_j(t_{wj}) + a_j)\hat{\Lambda}_j}{(\hat{\Lambda}_j t_{wj} + b_j)^2}$$

which is maximized when $\partial R/\partial t_{wj} = 0$. So:

$$\hat{\Lambda}_j\hat{a}'_j(t_{wj})(\hat{\Lambda}_j t_{wj} + b_j) - (\hat{\Lambda}_j\hat{a}_j(t_{wj}) + a_j)\hat{\Lambda}_j = 0$$

which when solved for $\hat{a}'_j$ results in:

$$\hat{a}'_j(t_{wj}) = \frac{\hat{\Lambda}_j\hat{a}_j(t_{wj}) + a_j}{\hat{\Lambda}_j t_{wj} + b_j} = \mathbb{U}$$

Since this is true for an arbitrary j, then we have the result that the optimal time in each task must be such that the instantaneous rate of gain in that task equals the average rate of gain across tasks. This result is equivalent to the marginal value theorem from Charnov (1976), and this derviation is similar to that from Pirolli (2007). Importantly, the original theorem only considers instantaneous reward, while our version allows the combination of instant reward and goals.

Influences derivation

We want to know how t^* changes with the parameters $= [a1, b1, B, L]$. Using the implicit function theorem the optimal allocations can be written as a function of θ, $\mathbf{t}^*(\theta) = \mathbf{t}^*(a_j, \mathbf{a_{-j}}, b_j, \mathbf{b_{-j}}, \ )$. Then we rewrite U as

$$\mathbb{U}(t_j(\theta), t^*_{-j}(\theta), \theta) = \frac{a_j(t^*_{w_j})}{t^*_{w_j} + L} + \frac{B}{t^*_{w_j} + L}$$

The fact that the time allocation problem is unconstrained means that the gradient of U with respect to the time allocations is zero.

$$\frac{\partial \mathbb{U}}{\partial \theta} = \frac{\partial \mathbb{U}(t^*_j, t^*_{-j}, \theta)}{\partial \theta} + \sum_i \frac{\partial \mathbb{U}}{\partial t_i} \frac{\partial t^*_i}{\partial \theta}$$

And thus

$$\frac{\partial \mathbb{U}}{\partial t_i} = 0$$

This envelope theorem result, together with the implicit value theorem, means we can compute the influence of changes in the optimal parameters through the direct effect on the mapping. Let

$$t^* = \phi(\theta)$$

if

$$f_i(t^*, \theta) = \frac{\partial \mathbb{U}}{\partial t_i} = 0$$

then

$$\frac{\partial \phi(\theta)}{\partial \theta} = -\nabla_\theta f\,(\theta, \phi(\theta))\,(\nabla_{t^*} f\,(\theta, \phi(\theta)))^{-1}$$

This is the rate of change of the gradient with respect to the Hessian, which is always 0 (given the above derivation), so we have:

$$\frac{\partial t^*}{\partial \theta} = \frac{\partial \phi(\theta)}{\partial \theta} = -\nabla_\theta f\,(\theta, \phi(\theta)) = -\frac{\partial^2 \mathbb{U}^2}{\partial t^*_{wj} \partial \theta}$$

Which is the result in equation 5.6.

Nonparametric Rate Equations

The main result is equivalent to an interacting rate process with constraints. Observations of in-task times and transitions can be characterized by a set of hazard functions that represent time-in-task data as an event stream $(k_1, t_1), \cdots, (k_N, t_N)$ of entering task type k_i at times $t_1 < \cdots < t_N$. Our theory induces a multi-variate point process. Each task has an underlying intensity rate $\lambda_i(t)$, which acts as an instantaneous urgency for task i, and $\vec{\lambda}$ denotes the set of these rates. In a multi-variate point process, we can predict the *time* and *type* of the next event. The next event's time t_i has density proportional to the sum of the intensities (because all events are competing) is given by $p(t_{i+1}|history) = \sum_k \lambda_k(t_{i+1}) \exp\left(\int_{t_i}^{t+1} \sum_k \lambda_k(t_{i+1})\right)$. Given this time t_{i+1}, the next event type is given by $p(K_{i+1} = k|history, t_{i+1}) = \lambda_k(t_{i+1}) / \sum_k \lambda_k(t_{i+1})$.

We can equate these rates with the instantaneous attainments associated with each task if we match key structural characteristics of the marginal rate theorem results. First, the marginal value theorem means that switches occur when the intensity for task k is equal to the the average intensity of other tasks, excluding task k. This means the probability of time till next task is the intensity should peak when $\lambda_k(t) = \sum_k \lambda_k(t_{i+1}) - \lambda_k(t)$ or $\lambda_k(t) = \sum_k \lambda_k(t_{i+1})/2$.

Interlude

The time allocation theory developed in Chapter 5 makes structural predictions about the relationship between background and foreground time use. In particular, they should be proportional to one another, due to their trade-off. We use these structural effects to make predictions of people's time use in two different data-sets: a mobile phone application switching dataset and the American Time Use Survey.

Chapter 6

Predicting contextual influences on time use via a rational model of time allocation

6.1 Introduction

Every day, millions of people pick up their mobile devices and log into their computers or tablets to engage in a productive activity, only to find themselves rapidly siphoned away by an alert, a game, or other distraction. Although technically voluntary, these switches can be extremely frustrating, are largely undesired and create problems of lost productivity, concentration and time. Users may even feel some loss of control. Increasingly, our devices themselves form background distractions that detract from our other leisure activities and socialization. Despite much attention, understanding and predicting the phenomena has proven difficult, with blame laid at the feet of the device for its addictive potential (Harris, 2017) or our brains, presumably too primitive for the modern world (Gazzaley & Rosen, 2016). Here we attempt to understand the phenomena by modeling how our brains naturally solve time allocation problems. We show that simple, rational principles for time allocation explain why our modern environment creates these challenges. Through this psychological computing approach (Bao, Gowda, Mahajan, & Choudhury, 2013), by understanding this natural switching phenomenon, we may leverage this understanding to help make better design decisions and uncover principles for guiding user experience.

One of the fundamental constraints on behavior is that not everything can be done at once. While this may be obvious, frustrating attempts at multitasking illustrate the principle: a primary goal, such as work or school, is interrupted by distractors, such as social media or emails (Jin & Dabbish, 2009; Gazzaley & Rosen, 2016) resulting in lost progress at diminished performance on both activities. Even if the distractor has a small time expenditure, the act of switching itself can impact performance. Research on task switching indicates that switching decreases performance in most tasks (Carrier, Rosen, Cheever, & Lim, 2015; Monsell, 2003), including app use (Leiva, Böhmer, Gehring, & Krüger, 2012), even when the switching is self-generated (Arrington & Logan, 2004). People display an inability to anticipate how task switches will reduce their performance on important tasks (Rosen, Mark Carrier, & Cheever, 2013), even though both external and internal distractions are common (Judd, 2015; Marulanda-Carter & Jackson, 2012). These distractions may be work-relevant, such as email, but not directly related to the task at hand (Marulanda-Carter & Jackson, 2012), while others might be internally driven by boredom or mind wandering (Kane, Kwapil, Mcvay, & Myin-germeys, 2007), or by alternative needs that are not fulfilled by the primary task (Wang & Tchernev, 2012). While there are benefits to engaging deeply in one task (Csikszentmihalyi & Lefevre, 1989), and people seem to be averse to switching (Kool, McGuire, Rosen, & Botvinick, 2010), people nevertheless regularly engage in switching behavior, even when it's identified as problematic. These phenomena suggest that the problem is not just a matter of user willpower, rather that users are responding to environmental pressures that encourage an internal competition between many different tasks.

People are readily aware of their own switching behavior, and have developed tools to help counteract it. In mobile phone use, existing tools to help users stay on task function either by restricting allowed applications or by providing some passive report on the foreground applications (e.g., Löchtefeld, Böhmer, and Ganev (2013)). These solutions are targeted to the hypothesized causes of switching behavior. Viewing apps as addictive leads to restricting or criminalizing their use. Viewing our brains as defective or limited leads to pathologizing the behavior in question and justifies efforts to treat or retrain our brains.

Here we show that switching behavior is rationally predicted by the structure of our environment — non-defective, non-addicted rational agents will do the same as other humans in the same kinds of environments. This provides an easier target for intervention. Ideally, we would like to restructure the user's environment to deincentivize undesired switching behavior. What is currently missing is a way to predict the incentive to switch. Using psychologically-grounded theory of the factors incentivizing app usage, we develop a foraging model for switching behavior that identifies the environmental and subjective forces that stimulate incentive to switch.

6.1.1 Predicting Time Use Behavior

Most of the effort in modeling phone application usage has focused on predicting the next application that will be used (Shin, Hong, & Dey, 2012; Böhmer, Hecht, Schöning, Krüger, & Bauer, 2011). These studies construct their prediction by applying machine learning approaches to the rich set of contextual features that are available on cell phones, like location, activity and time. Many useful user-experience improvements can be made by predicting the next app, including prefetching and context-aware menus (Shin, Hong, & Dey, 2012). Despite these successes, predicting the next app does not tell us if a user prematurely quit the previous app, nor the forces that affect app duration.

Time use studies in economics or sociology often treat time use as the result of people allocating a budget, with interest in how different demographic, economic, or sociological variables impact how people spend their time on various activities (Jara-Díaz & Rosales-Salas, 2017). This means that spontaneous switching is difficult to account for in a standard time budget framework.

People are subject to a wide range of immediate and external influences while allocating time; however, many of these influences are not able to be directly captured. Here, we provide an analysis based in foraging theory that shows that individuals should be influenced by their, possibly latent, *decision context*, which involves both a state of their needs, setting relative urgency, and the set of options psychologically available, or the contextual *decision set*.

This decision set is part of the environmental context for a person, and an estimate of this set and their natural activity durations in this context provides access to the latent urgency state of the user, that is, their baseline desire to switch activities. Direct measures of the decision set and durations would provide a way to empirically characterize an individual's salient environmental context and internal urgency state. Here we construct proxy measures of these quantities and show that they capture the predicted changes in switching behavior.

In general, activities in close temporal proximity to the current activities are the natural alternatives for a person's current activity choice, and their implicit availability and duration provide insight into the true decision context. Thus, we use the temporal proximity of activities to construct simple proxy measures for the decision set, but critically, we also incorporate recent activity *durations* into our proxy for decision context and show that durations have a strong contextual impact on switching.

Within mobile phone usage, this idea is similar to previous work using the *application chain* (the set of applications used sequentially before and after the app) as a contextual feature; however, decision context enlarges the feature to include durations and the concept of natural choice alternatives. Each member of the set of recent activities essentially acts as a candidate for switching. While this is a proxy choice, psychologically, it corresponds to people setting their decisions based on easily retrievable examples (the so-called availability heuristic (Tversky & Kahneman, 1973), rather than considering all choices.

The problem of choosing an activity and its duration is a time allocation problem directly analogous to the foraging problem for animals. A similar analogy has previously been used to model how users seek and consume information (Pirolli & Card, 1995; Pirolli, 2007). In an effort to bring understandings of human needs/psychology into the loop, we extend foraging theory's key results to more general human tasks, developing a time allocation model. This allows us to predict and verify contextual behavior in two different time usage datasets; a mobile phone usage dataset (Böhmer, Lander, & Krüger, 2013b) and time-use survey (United States., 2003).

6.2 Optimal time allocation theory

We draw upon foraging theory to construct a model of time allocation for *maximizing expected urgency rate*, $\mathbb{U}(t_{w1}, t_{w2}, \dots, t_{wn})$ where t_{wj} is the time allocated to the j^{th} task. Urgency represents the predictive decision quantity for the rate at which the decision maker can get value out of an activity, relative to alternatives. Urgency integrates all the factors that influence the chance of getting that value, typically task completion. Decision makers are rewarded if they allocate time in tasks that maximize the expected urgency $\mathbb{U}$, the average cumulative urgency experienced across tasks with time allocations t_{wj}. We show how to decompose this urgency rate into key psychological factors influenced by the environment: the availability of task alternatives and switch costs on the task outcomes. These factors modulate the attainment rates for tasks and make testable predictions about the impact of environment on time use behavior.

6.2.1 Modeling task value

We model the urgency of a task as the rate at which the user can get value out of it. The critical value for most activities are achieving events that indicate whether the task or components of the task are completed. The attainment is a temporal forecast of the probability of these events occurring given that the user engages in the task. The *attainment* $\alpha_j(t_{wj})$ within task j is a cumulative function which models the accumulation of the chance of these task-events occurring. We assume it is a monotonic increasing function of the time on task and also assume that $a_j(t_{wj})$ has a decreasing but non-negative derivative. This is equivalent to assuming that time on task decreases urgency, which prevents any one task from monopolizing time allocation. The attainment is modified by both switch costs and task availability.

First, we incorporate *switching costs* between tasks. Moving from one task to another often produces a cognitive cost associated with that switch, which can produce a functional loss in progress relative to baseline. We model this conditional cost using $\kappa_i(t_{wj}|t_{wi})$, a matrix associated with the loss of attainment for task j in time t_{wj} after having spent t_{wi} time in task i. In general, this function is a probability (i.e., takes values between 0 and 1) and increases with t_{wj} and decreases with t_{wi}; it is the total fraction of events kept given the possible loss due to the time in the alternative task.

The *availability* of the task j, represented by a (constant) rate λ_{ij}, depends on which task is being halted and acts as a switching cost. If a task is not available, you cannot switch to it.

We then construct an average attainment over the time allocated:
$A = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_{ij} T_B \alpha_j(t_{wj}) \kappa_j(t_{wj}|t_{wi})$. This produces an average urgency rate (within

tasks T_w and transitioning between them T_B).

$$\mathbb{U} = \frac{A}{T_w + T_B} \tag{6.1}$$

Note that this function represents an expectation across beliefs about task availability, task completability, and urgency signal availability, and includes dependencies between tasks. Despite the complexity, optimizing $\mathbb{U}$ produces a simple optimality condition, that is:

$$\frac{\partial \hat{A}_j}{\partial t_{wj}} = \mathbb{U}^* \tag{6.2}$$

meaning that the optimal time in a task is such that the instantaneous rate of gain in the task is equal to the average rate of gain at the optimal time allocations.

Equation 6.2 defines an optimal switching criteria: an agent should leave a task when the current instantaneous urgency gain ($\partial \hat{A}_j / \partial t_{wj}$) is equal to the average rate of gains in the environment ($\mathbb{U}^*$). Equation 6.2 represents a generalization of the Marginal Value Theorem from Charnov (1976) with minimal modifications to capture variables important in most human tasks.

6.2.2 The optimal task switching policy structure

The average rate of the environment (i.e., background rate) serves as a *threshold* for quitting a foreground task (see $\mathbb{U}^*$ in Figure 6.1). We can rewrite $\mathbb{U}$ by separating out the background influences of our time in a given task[1]:

$$\mathbb{U}(t_j, t_{-j}) = \frac{\alpha_j(t_{w_j})}{t_{w_j} + L} + \frac{B}{t_{w_j} + L} \tag{6.3}$$

The overall cost function can be decomposed into three distinct parts: factors that affect the foreground attainment α_j, factors that affect the average background attainment B, and

[1]We reduced the above equation by setting $B = (\sum_{i \neq j} \Lambda_i \alpha_i (t_{w_i^*}) / \Lambda_j)$ and $L = (\sum_{i \neq j} \Lambda_i t_{w_i^*} + T_B) / \Lambda_j$.

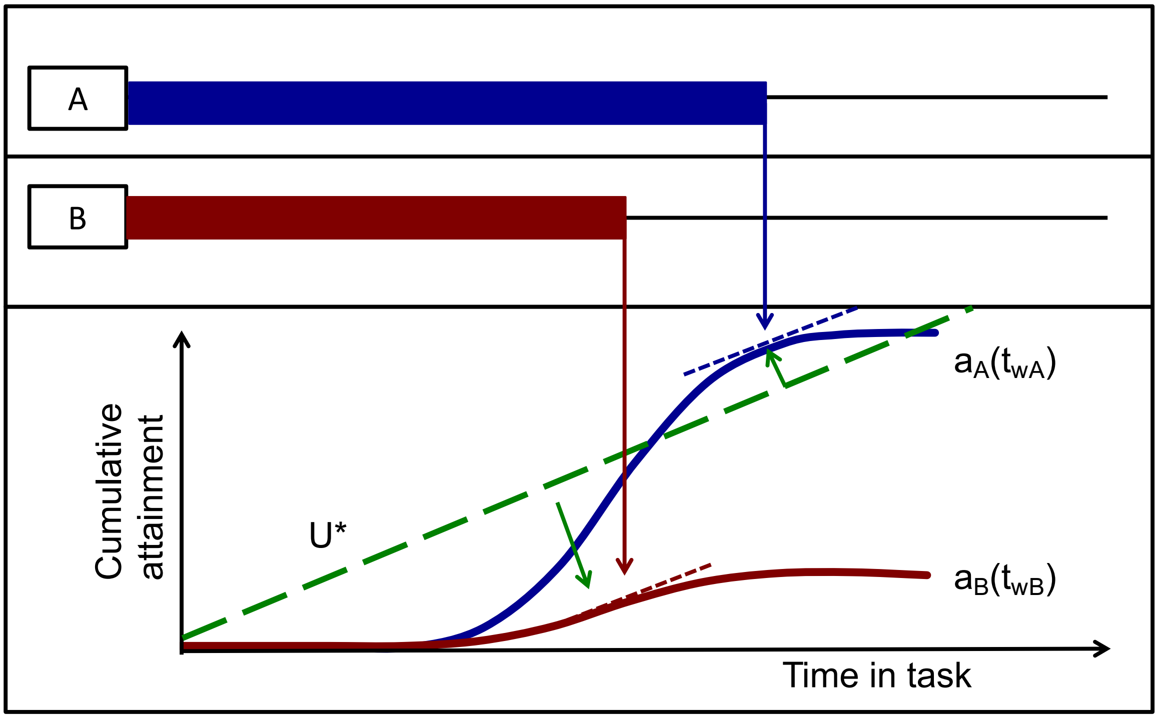

Figure 6.1: (Top) Optimal time in task as determined by marginal value theorem. (Bottom) The optimal switching point is when the rate of attainment drops below the average environment rate. This is visualized as the tangent lines whose slope is equal to the background attainment rate $\mathbb{U}^*$.

factors that affect the rate at which tasks can be engaged in or completed Λ_j (see Figure 6.2). Changes to each of these factors results in an effect on the duration of the foreground task. For this paper we are most interested in the effect produced by manipulating the background rate.

6.2.3 Environmental impacts on time on task

If the foreground or background urgency rates are markedly different, whichever has the higher urgency will receive all the time allocation. More generally, increasing background rate will *decrease* time in task, while increases in foreground rate will *increase* time in task.

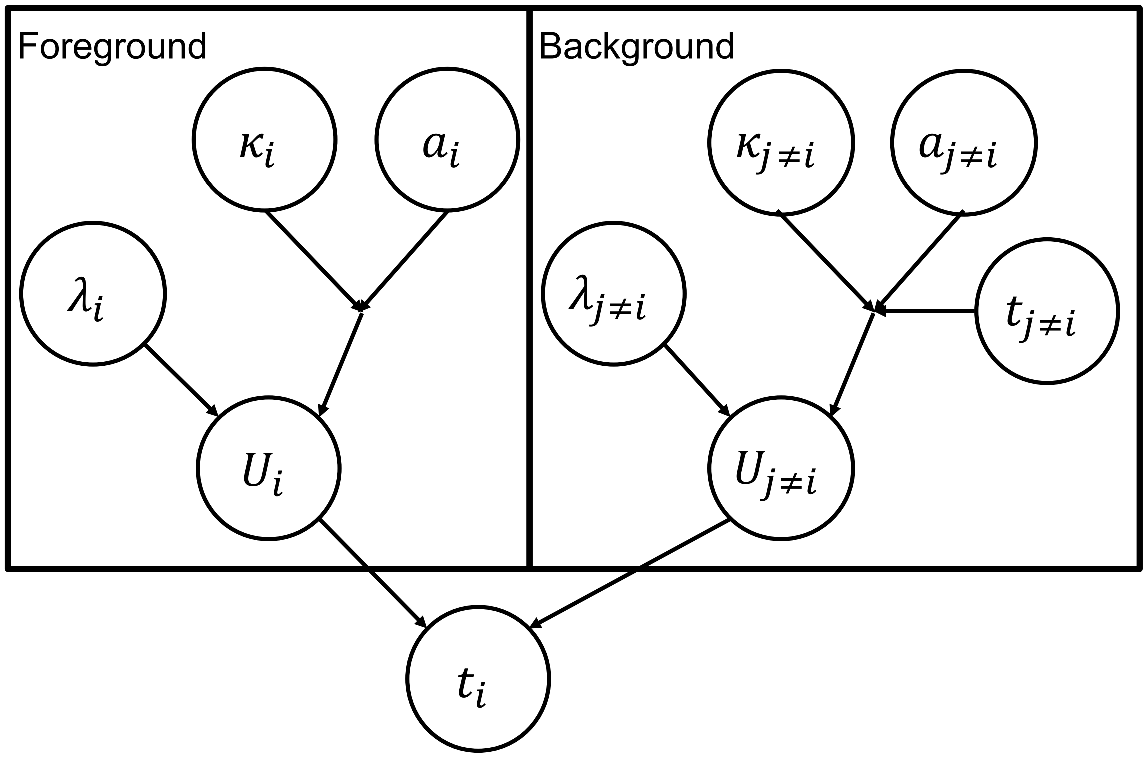

Figure 6.2: Direction of influences on time allocation in the foreground task, separating foreground and background factors. Attainment α_i, availability λ_i, and switch cost κ_i combine to influence the foreground rate $\mathbb{U}_i$ for task i. Similarly, background factors combine to form $\mathbb{U}_{i\neq j}$ for all background tasks, which combines with foreground $\mathbb{U}_i$ to produce within task time t_i.

Changes in relative urgency produce a variety of common app usage behaviors that are complex, but consonant with commonsense reasoning about the conditions that maximize productivity.

- *Distracting environment effect:* In the presence of a set of highly rewarding and quick-to-complete tasks, time in all tasks will decrease. In other words, you are more likely to be tempted to switch when there are lots of rewarding, low effort alternatives to switch to. In contrast, time in all tasks will increase with a background characterized by slow-to-complete, low urgency tasks. However, the environment is also devalued.

- *Interruptions and Notifications:* Events that change background rates affect decisions to stay in a task even without changes to the current task, including changes in decision context due to signals or reminders of background app availability. Notifications create a change in B; the increase in availability increases the background rate of tasks, shortening time in current task, and sometimes causing immediate switches.

- *App blocking "Nanny" effect:* More difficult transition between tasks (high Λ_j) incentivizes more time in task, up to a point. Switching times that are too long reduce time in task and in the limit, incentivize switching away from the device entirely.

- *Incentives for sticking:* The best scenario for allocating time to a long range project is: 1) high payoff (increased α_i), 2) competing projects are also long-range (increased B), and 3) switching times are high (increased T_B). It's clear our modern environment is poorly matched to these conditions.

Critical to all these effects is that the current activity's duration is being set relative to the context of what is provided by background alternatives, and the duration spent in background alternatives is a valid proxy for their relative urgencies. We carefully test this idea using two time use data sets.

6.3 Testing on Time Use datasets

To test the theory's prediction that the background rate is a significant factor in an individual's time in task, we constructed simple, measurable proxies for the background rate. The background is comprised of the decision set and its rates. Because these rates should affect foreground activity durations, we can capture the background rate using the decision context idea described in the introduction of this chapter.

There are many potential indicators of background rate that make use of different time windows and processing techniques (e.g., unsupervised learning via topic modeling to construct a mixture of categories for an app). While more elaborate definitions of decision context may be fruitful for prediction, here we focus on demonstrating that background rate per se as captured by proximally selected apps has the predicted impact on app switching.

The theory makes predictions based on the background and foreground rates, so it is useful to think about how durations relate to these rates. As the background rate increases, the time in a foreground app will decrease and vice versa. Therefore the durations of the background activities are inversely related to the background rate. By understanding this relation, we can use the durations of activities performed within close proximity as an indicator of the background rate.

Our theory predicts that environments with low background rates should also have a higher duration in the foreground task. Figure 6.3 illustrates this effect. If an activity's duration is dependent on its context, we would expect the duration to increase when it is used in contexts with lower background rates.

Given the expected background time (average time across background activities) for a given context $\mathbf{E}(t_{\text{background}}|c) = T_{background}[c]$, we expect that:

$$T_{foreground}[c, activity] \propto T_{background}[c] \tag{6.4}$$

for activity identifier ($act \in Acts$) and context c. As the expected background average time increases for a given context, the time in a foreground activity should also increase.

We can show that the relationship in equation 6.4 holds by performing a linear regression between the foreground and background times. For each activity and background context

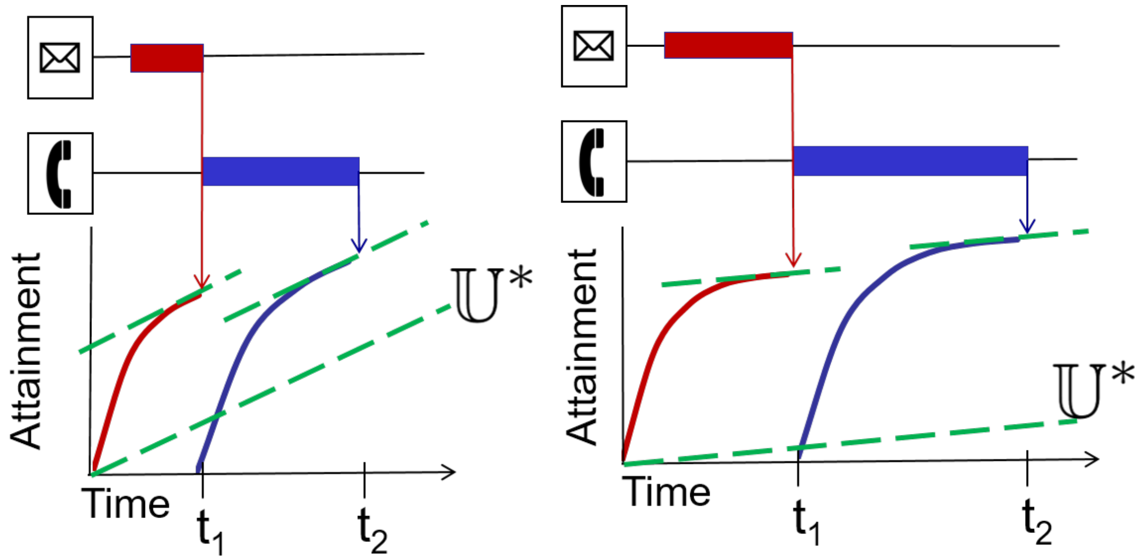

Figure 6.3: Marginal value theorem allows us to make measurable predictions to task allocation based on context changes. The effect of decreasing the background rate $\mathbb{U}*$ *increases* the time in task. (Left) High background rate, low duration. (Right) Low background rate, high time in task.

we construct, we collect the following sets of data:

$$X = \{T_{background}[c] | act_i = act\} \tag{6.5}$$

$$Y = \{t_i | act_i = act)\} \tag{6.6}$$

We then perform standard linear least-squares regression $Y = \beta X$, and record both the slope β and the two-sided p-value for a hypothesis test whose null hypothesis is that the slope is zero, using a Wald Test with t-distribution of the test statistic. Based on the presented theory, we would expect a positive correlation between the duration of an activity and the duration of the background alternatives.

To show that our choice of decision context proxy does not need to be exact, we try two simple measures: using the mean duration of the previous and next activity where background duration $= \frac{1}{2}(t_{i-1} + t_{i+1})$ and using only the duration of the previous activity where background duration $= t_{i-1}$. The latter construction is more interesting for online predictions of background rate.

6.3.1 Smartphone App Use Dataset

Smartphones have grown to support a variety of work and leisure activities and have become a pervasive fixture of modern life. The availability of a diverse set of tasks makes cellphone usage an ideal natural testing environment for evaluating a theory of task scheduling. We analyze a dataset of application visitation behavior on smart phones from Böhmer, Lander, and Krüger (2013b). This dataset has been previously used to reveal application usage patterns (Böhmer, Ganev, & Krüger, 2013a; Leiva, Böhmer, Gehring, & Krüger, 2012; Böhmer, Hecht, Schöning, Krüger, & Bauer, 2011; Parate, Böhmer, Chu, Ganesan, & Marlin, 2013).

Our dataset has over 58 million instances of application use from 7731 users, with 88,150 unique applications (see Figure 6.4). The mobile dataset includes which application is active at a given time for a given individual's phone (see Figure 6.5 for relevant attributes). Notably, this dataset is focused on measuring the "primary" use of the phone rather than all active applications; multitasking is not assessed within this dataset (e.g., listening to music while browsing the internet). This focus on primary applications produces a task-switching structure as shown in Figure 6.6.

We treat individual applications as unique activities for this analysis. Our contexts are therefore defined based on the "nearby" applications, primarily those just before and after the use of a given one. This produces a large possible set of contexts $(88,150^2)$, which must be filtered to produce interesting contexts. We only treat an application as being used in a context if it had been in a context more than 40 times, and we only use contexts that have at least 5 applications in that context.

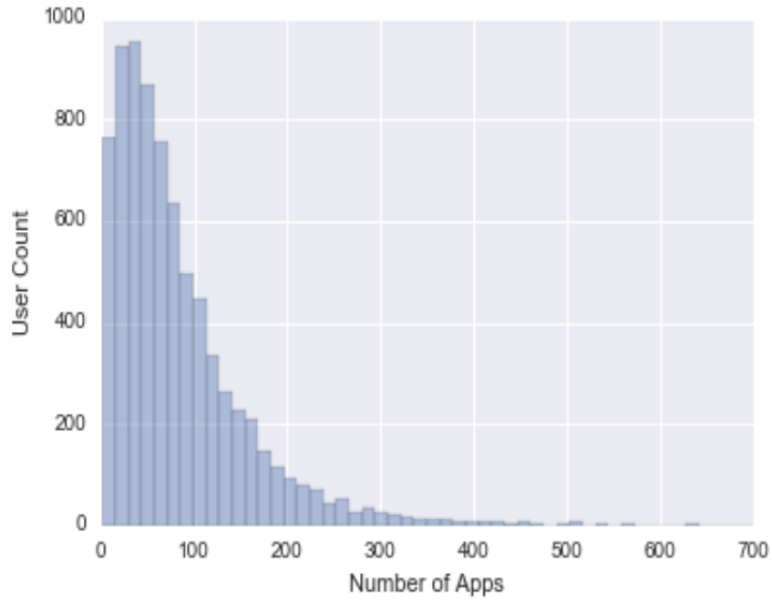

Figure 6.4: Histogram of number of applications per user.

Attribute	Aescription
device_id	unique identifier for the user
app_id	md5 hash to identify app
start_time	utc timestamp
run_time	millisecond runtime
session_start_time	utc timestamp for session

Figure 6.5: Attributes in the smartphone app usage dataset.

Figure 6.6: Example of application switching within the dataset. This user transitions from one application to another, after some period of time. We construct our notion of context using this app switching pattern, by taking "nearby" applications as our context.

6.3.2 American Time Use Survey Dataset

To verify that our results can be extended beyond the particular dataset we chose, we applied the same analysis to the American Time Use (ATUS) dataset (United States., 2003). The ATUS dataset is a time diary-based survey conducted by the United States Bureau of Labor Statistics (BLS) to assess the activities most citizens spend their daily time on. In the ATUS, individuals and households report what primary activity they performed in each 15 minute increment throughout a single conterminous 24 hours (starting at 4am and ending at 4am). The activities are coded based on a standardized schema, designed to ensure uniformity across individuals and reliability of coding. While some time use surveys measure multitasking, these activity codings focus on the primary activity at the time (e.g., eating while socializing would be encoded as one or the other depending on primary purpose). Activities are coded in a three-layer hierarchy (e.g., "household activities," "housework," "laundry"), with options for ambiguous codings (e.g., housework that does not conform to other definitions). An example of a single day for an individual can be seen in Figure 6.9.

The ATUS is collected with the goal of associating time usage on different activities with various demographically-, economically-, or sociologically relevant features within the United States (Hamermesh, Frazis, & Stewart, 2005). For example, time-use studies in the social sciences investigate the impact of "unpaid labor"' on the overall economy, on changes in behavior over time (e.g., the rise of digital devices), or gender or economic differences in activities (e.g., labor/leisure tradeoffs) (Cornwell, Gershuny, & Sullivan, 2019). The ATUS dataset also includes a large set of survey and demographic information on each individual and household recorded; however, here we focus on just the activities for our current analysis (see Figure 6.8 for relevant attributes).

The ATUS dataset is freely available on the BLS website[2]. General information, including data collection procedures and example publications, can be found there as well. We used the multi-year dataset from 2003 to 2018; this includes 201,151 unique individuals and 426 unique activities defined based on a standardized coding scheme (Hamermesh, Frazis, & Stewart, 2005), resulting in 3.9 million activity instances. To compare with the app-usage dataset, we treat phone applications and activities as roughly synonymous for the purpose of our analysis.

Of critical importance is that the activities considered in the ATUS dataset include activities which are heavily time constrained, such as work or medical appointments, which are not as subject to the spontaneous time allocation that our theory focuses on and which require integrating external time constrains differently. For now, we focus on activities which can be freely scheduled rather than those which are constrained. We follow Flood, Hill, and Genadek (2018) in a separation of activities into four categories: committed time, contracted time, necessary time, and free time. Contracted activities are those such as education and work that are heavily regulated and synchronized across a society; committed activities are generally those of unpaid labor that result from prior commitments, such as house and care work (e.g., parental care). Necessary activities generally are basic biological need-fulfilling activities, such as sleeping and eating, that must be engaged in during a day. Free time is essentially defined as any time not under the other categories, but is primarily

[2] https://www.bls.gov/tus/home.htm

leisure activities such as sports and socializing.

Here we focus our analysis on predicting free-time activities, resulting in a smaller set of 180 unique activities with ~ 1 million activity instances. However, we allow the other activities to define our background context (see Figure 6.9). We focus only on contexts with more than 50 samples below.

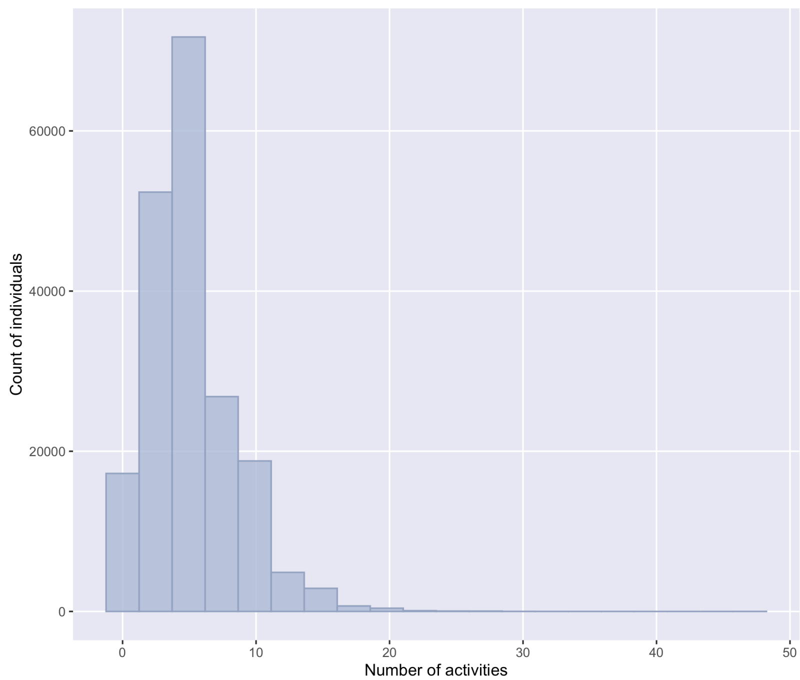

Figure 6.7: Histogram of "free time" activities per individual.

Attribute	Description
TUCASEID	unique case id for the individual
TRCODEP	six-digit activity code
TUACTDUR	activity duration

Figure 6.8: Attributes in the ATUS dataset. Note that only a single individual's daily activities are measured per household, so TUCASEID refers to both a given household and individual. For our purposes, we simply use it as an individual's identifier.

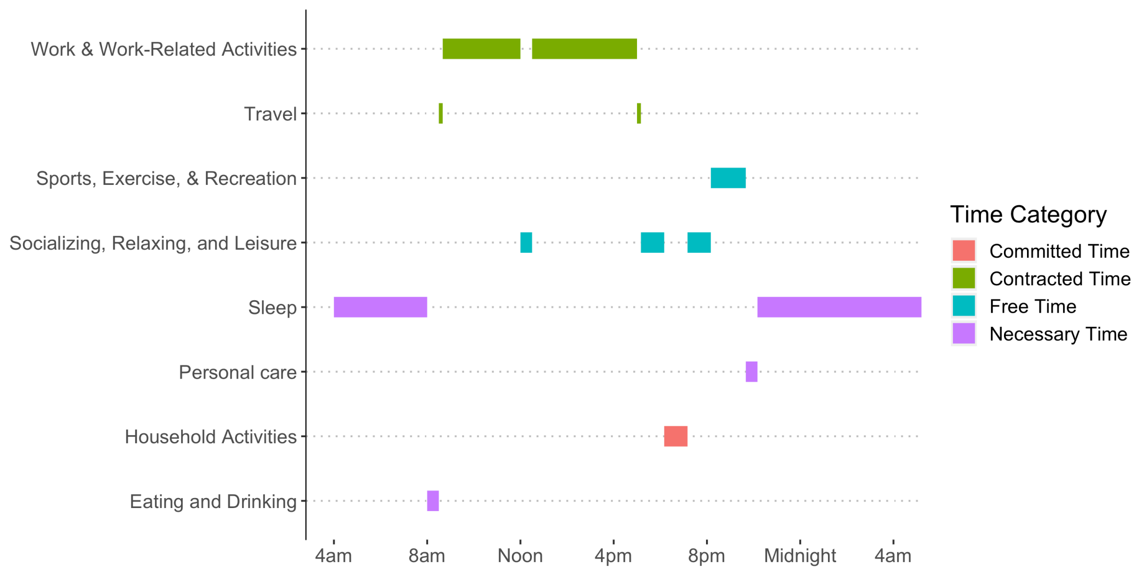

Figure 6.9: Example of a single day from a respondent in the ATUS, represented as a Gantt chart. Here we show which activities they engage in at different times. Activity labeling is based on the highest level, except for sleep (which is part of "personal care"). The time categories are defined in Flood, Hill, and Genadek (2018). Note that for our analysis we focus on predicting activities in "free time," but allow other categories to define the background. In this example, the socializing that occurs around noon is surrounded by work activities, which will define the local background for that socializing activity.

6.3.3 Results for smartphone app usage

For both constructions of our decision context proxy, the majority of the regressions had significant ($p<.05$) positive slopes. Figure 6.10 shows the distribution of these slopes. The context that made use of the previous and next app durations yielded 530/618 ($\approx .86$) significant slopes with 525/530 ($\approx .99$) of these slopes being positive. The context that made use of only the previous app duration yielded 511/618($\approx .83$) significant slopes with 506/511($\approx .99$) of these slopes being positive. This difference is reasonable because the latter makes use of less context information, and is therefore more sensitive to noise from the previous app's duration.

To ensure that the results are not tainted by a chance sampling of individual differences, we reran the the analysis, except this time performing each regression on a per user basis. We ignore (user,app) regressions where the number of instances of the app is fewer than 30 for the user.

$$X = \{\text{background duration}_i | app_i = app, user\} \tag{6.7}$$

$$Y = \{t_i | app_i = app, user)\} \tag{6.8}$$

26458 of the 100219 regressions (≈.26) had a significant slope, due to each user having less data on any particular app. For the (user,app)s whose slopes are significant, the bottom of Figure 6.10 shows that the slopes are still overwhelmingly positive: 25760/26458 (≈.97).

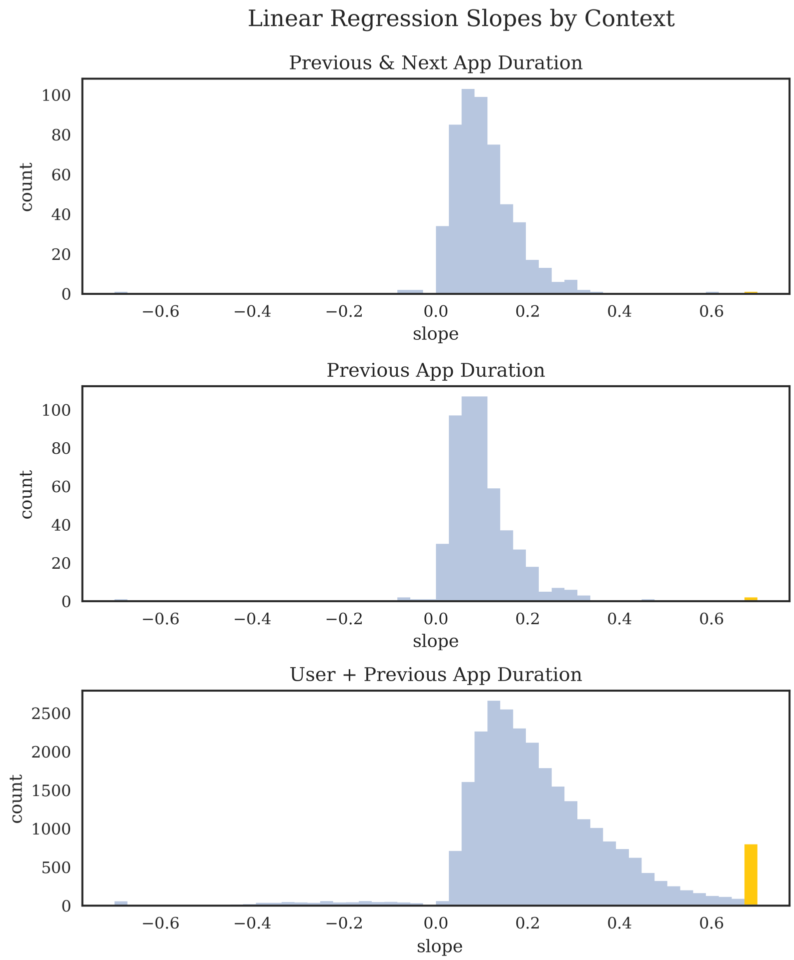

Figure 6.10: The distribution of significant slopes after performing a linear regression on the foreground and background durations for different constructions of context: (top) Using the previous and next app durations. (middle) Using only the previous app's duration. (Bottom) Results of performing analysis on a per-user basis instead of across users to ensure to ensure individual differences cannot account for the effect. Notice, almost all significant slopes are positive which indicates the time in an app is dependant on the background rate.

6.3.4 Results for ATUS dataset

For the ATUS dataset we found very similar results for the decision contexts. For both constructions of our decision context proxy, the majority of regressions with significant slopes were positive (same significance level, $p < 0.05$). Figure 6.11 shows the distributions of these slopes. The context with previous app durations yielded 610/2321 (≈.26) significance slopes with 588/610 (≈.96) being positive. The context using both the previous and next app durations yielded 1247/5724 (≈.22) significant slopes with 1193/1248 (≈.96). Due to sparseness of samples, we cannot perform a per-individual version of this analysis for the ATUS dataset (no user has more than 20 instances of an activity, across contexts).

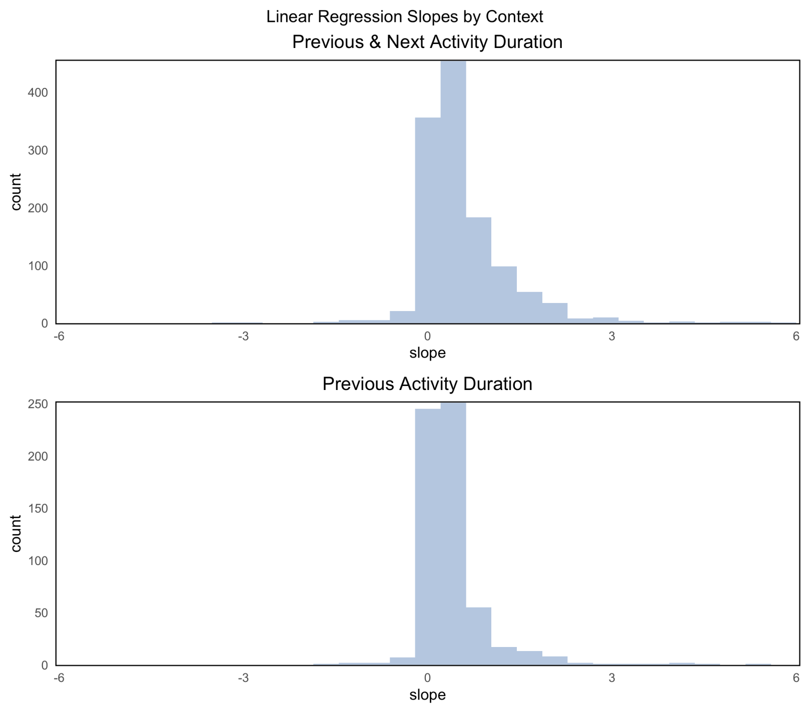

Figure 6.11: The distribution of significant slopes after performing a linear regression on the foreground and background durations for different constructions of context: (top) Using the previous and next activity durations. (middle) Using only the previous activity's duration. Notice, almost all significant slopes are positive, which indicates that the time in an activity is dependant on the background rate.

6.4 Conclusion

We developed a foraging theory model of app switching that identifies key structural factors that impact time allocation. We identify the user's *decision context* as the critical modulator of activity switching that can be impacted by environmental factors and show that simple proxy measures reliably predict changes in activity durations.

We interpret the positiveness of significant slopes as verification of our predictions that the background rate has a direct and measurable impact on the foreground time in activities. To further illustrate this importance, the mean slope of 0.10 for the previous+next construction for phone apps represents a 10 percent increase in foreground duration relative to the background duration. Consider two contexts A and B, where context A has an average background duration of 10 seconds and B has an average background duration of 20 seconds. We could expect an app that usually takes 10 seconds in context A to take 11 seconds in context B due to the lower background rate, i.e., the lower incentive to switch out of the task.

Understanding human time allocation requires large time use data sets. The ATUS data set, for example, is large yet still smaller than the mobile phone app usage data set (by 100 orders of magnitude). The difference in results between the mobile use data set are likely due to sparseness of appropriate measures. The ATUS has both a smaller set of activities and fewer samples per activity; the structural relationships we are interested in require measuring the same activity for an individual across multiple contexts multiple times. Despite this difficulty, we still replicate our analysis across these data, showing the robustness of our results.

Our current focus has been on analyzing free time, in the context of the time-use survey. However, constrained and controlled time are important parts of contemporary time use that a scheduling theory must include. Incorporating the impact of forced time or deadlines into our theory can allow us to extend these results to other qualitatively different types of time allocations. Such an extension would be highly relevant to researchers studying how various social forces can control individual time use (Gerstel & Clawson, 2018).

More sophisticated proxies for decision context can also be developed, possibly incorporating the co-occurring app choices, logs, interrupts, and user preferences, as well as additional telemetry data that holds promise to improve app switching prediction. Future work with time use data sets can incorporate other demographic or survey results (e.g., different "modules" such as the well being module (Lee, Hofferth, Flood, & Fisher, 2016) that provide insight into their decision context. For example, the simple difference between a workday and the weekend can provide a needed context shift that is not currently incorporated.

Recent time use studies have identified an overall structure to how people allocate time in their day, showing that there are prototypical behavioral sequences that most people follow (Vagni & Cornwell, 2018; Flood, Hill, & Genadek, 2018). While these patterns could be used to provide a decision context to extend our results, future work should also investigate to what degree these patterns can be expected due to optimal time allocation across a population. It is likely that external constraints would need to be incorporated, along with a population-level analysis.

Our work here demonstrates that extending foraging theory to human time use provides predictable results. The core feature we exploit is the natural trade-off between alternative activities, which produces a background context that impacts immediate time allocation. This also demonstrates that task switching can be rationally predicted from contextual features, suggesting that solutions to these problems should take the form of environmental rather than individual change.

Chapter 7

Discussion and Conclusion

In this thesis, we have looked at the problems motivation must solve, in particular, time and resource allocation. We have provided a theoretical framework to understand how motivation interacts with other cognitive and emotional processes. In Chapter 2 we explored the phenomenon of motivated engagement and constructed a theory of scheduling that specifies computational principles that motivation likely follows. In particular, we determined that the priority of goals for engagement requires monitoring various environmental and task cues for priority. We have investigated how decision making is impacted by motivational and emotional systems, and how time allocation is a core problem that motivation must solve. We then provide evidence for how these meta-cognitive systems interact and how the structural implications of time allocation can be revealed in human time use.

We have discussed how to extract metacognitive signatures of both environmental and task quality. We looked at examples of how this monitoring process occurs that might set decision parameters. In Chapter 3 we developed a method to extract decision-relevant modulatory processes. By using projection to latent structures regression[1], combined with a decision-process model (DDM), we found a latent space that sets decision parameters in a coordinated way. In this instance, the environmental variable was straightforward, likely a cue for environmental danger (Mobbs et al., 2015, FEB). In Chapter 4, we extended the

[1]The co-creator of partial least squares, Svante Wold, preferred this name (Wold, Sjöström, & Eriksson, 2001)

decision-process model to make use of metacognitive task monitoring. Confidence, we showed, can act as an efferent forecast of information reliability that the decision process can use to dynamically change decision parameters. In both cases, the standard drift diffusion decision model had to be extended to incorporate hierarchical monitoring, which has fundamentally limited the DDM as a framework to study resource allocation and optimal stopping in humans.

In Chapter 5, we focused on developing a time allocation theory for human engagement by extending the standard optimal foraging theory "patch" model. We extended it to account for tasks more common in human behavior, specifically tasks with completion logic and those incorporating only intrinsic motives. We showed how these extensions can be applied across human task switching and can predict a wide range of relevant phenomena. An immediate next step for the time-allocation theory developed in Chapter 5 is to apply it to the study of human time use. We used our foraging theory of engagement to predict human time allocation in two time use data sets in Chapter 6, showing that the structural predictions the theory makes can be practically used.

One of the major goals of this dissertation is to provide clarity on what motivation is. *Motivation* is an unfortunately broad term. As previously mentioned, its most generic version refers to the underlying causes of human behavior. However, as emphasized by Tinbergen (1963), Marr (1982), among others (Anderson, 1990), there are multiple "why" questions that can be asked about human behavior, depending on the level of granularity, timescale, or abstraction. To explain all of human behavior, one might, for instance, take a control perspective (Cisek, 1999), as we have done here, where the cognitive and motivational processes our brain performs are part of that control architecture.

Within cognitive science, we often focus on understanding the human mind by studying how people behave within a task and form models of how people solve those task. The use of optimality theory greatly aids this process, as human behavior can often be derived based on a constrained optimization towards some cost minimization. Debates over terminology aside, optimality theory essentially allows the understanding of a dynamical system (such as behavior) through the specification of invariates (goals and constraints). This can result in a decision process model or computational architecture, allowing us to perform simulations or data fits.

This process of separating human behavior into tasks and understanding the computational processes needed to solve them has driven most research in the cognitive sciences. Arguably, the use of compositionality justifies this process. Compositional problems allow the problem space to be split into subcomponents, each of which are simpler to solve and can be solved without dealing with the other problems. Hierarchical systems likely evolve because of this property, producing a system that can solve complex problems while retaining simplicity by splitting the solution space for subcomponents to solve (Herbert et al., 1962). The various ways of separating unobserved psychological components (e.g., memory, emotion, attention, motivation) are also partially justified from this perspective, in that each component corresponds to distinct but interrelated information processing for the purpose of solving decision problems.

Motivation, from this instrumental view, arises due to the distinct problem of having to orchestrate resources across tasks. In particular, time allocation arises, due to mutually exclusive tasks requiring some priority computation that decides task engagement. In this dissertation we have taken the view that the subjectively experienced motivational impetus we feel (or don't feel), is due to the task priority computation our brain performs (often implicitly). This provides a very concrete perspective on motivation, synthesizing past research.

A core difficulty we have faced is due to motivation's various uses in the field, especially historically. Danziger (1997) traces the history of various psychological concepts as the field matured between the 19th and 20th centuries. Motivation developed as a concept that used biological and engineering ideas (e.g., drive and energy) as scientific backing to applied domains in industry and marketing. The distinction that motivation allowed was to explain how a worker might have poor performance despite high skill, as measured by the then-recently-developed psychometric technology of intelligence tests. What motivation research then provided for this burgeoning field was scientific authority on a rather broad array of applied topics. The fact that "motivation" was such a broad concept is actually significant to this development.

The various motivation theories that resulted often occurred because the phenomenon studied by these researchers varied so much. One of the major goals in this dissertation was to be clear about what particular phenomena *motivation* refers to in the context of engagement, what problem it solves, what computational processes are necessary to solve it, and the resulting implications on behavior. This does not exclude other definitions of engagement or motivation; multiple definitions that share conceptual tools can be useful, as long as clarity is maintained on the domain of application. However, we have provided a clear through-line on a particular view on motivation, connecting it with many of the common everyday phenomena motivation reflects.

This dissertation uses more sophisticated data-science techniques combined with psychological modeling. This technique of using structure derived from optimality theory (in this instance, optimal scheduling), can constrain machine learning techniques that find structure in data.

For example, our use of projection to latent structures in Chapter 3 also can be applied to various data sets where a large set of task-relevant variables are collected, but their relation to the decision process model is unknown. For example, for subject measures of executive function, rather than specifying a particular structure, as confirmatory factor analysis does, we can allow the structure to emerge based on their particular impact on any given task we wish to study. Examining how executive functions impact a video game could be performed by creating a decision-process model of the game and augmenting it with a more general modulatory system. This could similarly allow us to investigate how emotional states interact with cognitive processes by interfacing biometric measures with an agent model (Silver et al., 2016).

Another interesting extension of this work can be applied to gamification of learning environments. Gamification refers to the hope of using video-game-style design as a way of encouraging students to spend more time and to more deeply engage in learning. However, gamification work has met with mixed success (Dickey & Meier, 2005; Ryan, Rigby, & Przybylski, 2006). Given Chapter 5, time spent learning will trade off with other tasks a student is engaged in. This means many environmental factors are relevant to understand time use. We also need to understand what cues students use to determine task priority, as discussed in Chapter 2, in particular, task progress and intrinsic learning cues.

By using a model of student performance, we can create measures of empowerment and learnability in knowledge space (as in deep knowledge tracing àla Piech et al., 2015). Using these measures, we can then apply time allocation theory to predict when students will quit a learning task. For instance, students will likely quit learning tasks when they think information saturates (either due to their own skill or the material). From there, the question is what possible interventions can be applied that might change their expectations of the information available or change what learning task they are performing to keep them engaged. This potential future research can build on the results in this dissertation. However, questions about education are inherently political (*Weizenbaum examines computers and society*, n.d.) and require us to consider the political implications of controlling students' motivations.

A common concern in motivation is control over motivation; common questions include "How can I make students do their assignments," "How do I make my child behave," and "How can I stop being distracted by irrelevant tasks?" Most of these questions are either innocuous or meant in a positive way. Understandably, individuals want to be able to control their own behavior. A common feature of disorders such as attention deficit hyperactivity (ADHD) or depression is the feeling of being unable to control your own motivation[2], which can be both incredibly debilitating and frustrating for those experiencing it.

However, these ostensibly scientific discussions can hide ideological assumptions (Walsh, 2013). What is a productive or "good" use of time is both historically and culturally contingent (Thompson, 1967; Glennie & Thrift, 1996; White, Valk, & Dialmy, 2011) and often decided by those in power. An important discussion within the neurodiversity movement, for example, is whether these frustrations and "failures" are due to personal pathologies or structural issues within our culture and society. Is a given problem due to an individual's neurological makeup, or is it that much of our society only values people as producers of labor?

While practical discussions of motivation do refer to control over ourselves, they also refer to control over others. While a greater understanding of motivation can be used positively, such as for the improvement of education or for self-improvement, it can also be used by employers to shape workers or by the state to control prisoners. Current debates over the use of machine learning in large-scale surveillance systems directly relate to these issues. As researchers, it is our responsibility to consider these decidedly political implications and seriously think through what values we have and how those are manifested in our work.

[2]See `https://gekk.info/articles/adhd.html` and (Brosh, 2013) for personal descriptions of each.

[T]here are those who hope that the good of a better understanding of man and society which is offered by this new field of work may anticipate and outweigh the incidental contribution we are making to the concentration of power (which is always concentrated, by its very conditions of existence, in the hands of the most unscrupulous). I write in 1947, and I am compelled to say that it is a very slight hope.

(Wiener, 1961)

“I wish it need not have happened in my time,” said Frodo.

”So do I,” said Gandalf, ”and so do all who live to see such times. But that is not for them to decide. All we have to decide is what to do with the time that is given us.”

– J.R.R Tolkien, *The Lord of the Rings, The Fellowship of the Ring*

References

/Remake FAQ for League of Legends (2016). URL: https://support.riotgames.com/hc/en-us/articles/208754686%E2%80%93Remake-FAQ.

A.H. Maslow (1943). "A theory of human motivation". In: *Psychological Review* 50.4, pp. 370–396. ISSN: 0033-295X. DOI: 10.1037/h0054346. URL: http://psychclassics.yorku.ca/Maslow/motivation.htm.

A/S, iMotions (n.d.). *iMotions Biometric Research Platform*. Copenhagen, Denmark.

Aalen, Odd O. et al. (2008). *Survival and Event History Analysis: A Process Point of View*. 540 pp. ISBN: 0-387-68560-X. DOI: 10.1007/978-0-387-68560-1. URL: https://books.google.com/books?hl=en&lr=&id=2toprArSUMAC&pgis=1.

Aarts, Henk and Andrew Elliot (2012). *Goal-directed behavior*. Ed. by Taylor & Francis. Taylor & Francis. ISBN: 978-0-203-86966-6. DOI: 10.4324/9780203869666.

Achermann, P and A A Borbely (2003). "Mathematical models of sleep regulation". In: *Front Biosci* 8, s683–93. ISSN: 10939946. DOI: 10.2741/1064. URL: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12700054.

Acuña, Daniel E. and Paul Schrater (2010). "Structure learning in human sequential decision-making". In: *PLoS Computational Biology*. ISSN: 1553734X. DOI: 10.1371/journal.pcbi.1001003.

Addicott, Merideth A et al. (2017). "A primer on foraging and the explore/exploit trade-off for psychiatry research". In: *Neuropsychopharmacology* 42.10, p. 1931.

Adler, Rachel F and Raquel Benbunan-Fich (2013). "Self-interruptions in discretionary multitasking". In: *Computers in Human Behavior* 29.4, pp. 1441–1449.

Adler, William T (2018). "Computational mechanisms underlying human confidence reports". In: (January).

Adler, William T. and Wei Ji Ma (2018). "Limitations of proposed signatures of Bayesian confidence". In: *Neural computation*. ISSN: 0010-4655. DOI: 10.1101/218222. URL: https://www.biorxiv.org/content/early/2017/11/13/218222.

Aggarwal, Jk and Ms Ryoo (2011). "Human activity analysis: A review". In: *ACM Computing Surveys (CSUR)* 43.3, 16:1–16:43. ISSN: 03600300. DOI: 10.1145/1922649.1922653. URL: http://dl.acm.org/citation.cfm?id=1922653.

Alexander, B K et al. (1978). "The effect of housing and gender on morphine self-administration in rats." In: *Psychopharmacology* 58, pp. 175–179. ISSN: 0033-3158. DOI: 10.1007/BF00426903.

Anderson, David J and Pietro Perona (2014). "Toward a science of computational ethology". In: *Neuron* 84.1, pp. 18–31.

Anderson, J. R. (1990). *The adaptive character of thought*. Vol. 104. 276 pp. ISBN: 0-8058-0419-6. DOI: 10.4324/9780203771730.

Anderson, John R (1993). "Problem solving and learning." In: *American Psychologist* 48.1, p. 35.

Andersson, Patrik et al. (2012). "Real-time decoding of the direction of covert visuospatial attention". In: *Journal of Neural Engineering* 9.4, p. 045004. ISSN: 1741-2560. DOI: 10.1088/1741-2560/9/4/045004.

Arrington, Catherine M and Gordon D Logan (Sept. 2004). "The cost of a voluntary task switch." In: *Psychological science* 15.9, pp. 610–5. ISSN: 0956-7976. DOI: 10.1111/j.0956-7976.2004.00728.x. URL: http://www.ncbi.nlm.nih.gov/pubmed/15327632.

Ås, Dagfinn (1978). "Studies of Time-Use: Problems and Prospects". In: *Acta Sociologica*. ISSN: 15023869. DOI: 10.1177/000169937802100403.

Averbeck, Bruno B (2015). "Theory of choice in bandit, information sampling and foraging tasks". In: *PLoS computational biology* 11.3, e1004164.

Bacon, Pierre-Luc et al. (2017). "The option-critic architecture". In: *Thirty-First AAAI Conference on Artificial Intelligence*.

Baddeley, Alan (2000). "The episodic buffer: A new component of working memory?" In: *Trends in Cognitive Sciences* 4.11, pp. 417–423. ISSN: 13646613. DOI: 10.1016/S1364-6613(00)01538-2.

Badre, David (2008). "Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes". In: *Trends in cognitive sciences* 12.5, pp. 193–200.

Baker, Ryan S.J.d. et al. (Apr. 2010). "Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments". In: *International Journal of Human-Computer Studies* 68.4, pp. 223–241. ISSN: 10715819. DOI: 10.1016/j.ijhcs.2009.12.003.

Ballard, Timothy et al. (2018). "On the pursuit of multiple goals with different deadlines". In: *Journal of Applied Psychology* 103.11, pp. 1242–1264.

Bang, Dan et al. (2014). "What failure in collective decision-making tells us about metacognition collective failure and metacognition". In: *The Cognitive Neuroscience of Metacognition* 9783642451, pp. 189–221. ISSN: 14712970. DOI: 10.1007/978-3-642-45190-4_9.

Banis, Stella and Monicque M. Lorist (2012). "Acute noise stress impairs feedback processing". In: *Biological Psychology*. ISSN: 03010511. DOI: 10.1016/j.biopsycho.2012.06.009.

Bao, Xuan et al. (2013). "The Case for Psychological Computing". In: *Proceedings of the 14th Workshop on Mobile Computing Systems and Applications*. HotMobile '13. New York, NY, USA: ACM, 6:1–6:6. ISBN: 978-1-4503-1421-3. DOI: 10.1145/2444776.2444785. URL: http://doi.acm.org/10.1145/2444776.2444785.

Barabasi, Albert-Laszlo (2005). "The Origins of Bursts and Heavy Tails in Human Dynamics". In: *Nature* 435.7039, pp. 1–15. ISSN: 1476-4687. DOI: 10.1038/nature03459. URL: http://www.nature.com/nature/journal/v435/n7039/abs/nature03459.html.

Basten, U. et al. (2010). "How the brain integrates costs and benefits during decision making". In: *Proceedings of the National Academy of Sciences*. ISSN: 0027-8424. DOI: 10.1073/pnas.0908104107.

Baumeister, Roy F (2002). "Yielding to temptation: Self-control failure, impulsive purchasing, and consumer behavior". In: *Journal of consumer Research* 28.4, pp. 670–676.

Baumeister, Roy F. and Todd F. Heatherton (1996). "Self-Regulation Failure: An Overview". In: *Psychological Inquiry* 7.1, pp. 1–15. ISSN: 1047-840X. DOI: 10.1207/s15327965pli0701.

Baumeister, Roy F. and Kathleen D. Vohs (Nov. 2007). "Self-Regulation, Ego Depletion, and Motivation". In: *Social and Personality Psychology Compass* 1.1, pp. 115–128. ISSN: 1751-9004. DOI: 10.1111/j.1751-9004.2007.00001.x. URL: http://doi.wiley.com/10.1111/j.1751-9004.2007.00001.x.

Becker, Gary S (1965). "A Theory of the Allocation of Time". In: *The economic journal*, pp. 493–517.

Beersma, Domien (1998). "Models of human sleep regulation". In: *Sleep Med Rev* 2.1, pp. 31–43. ISSN: 10870792. DOI: 10.1016/S1087-0792(98)90052-1.

Berkay, Dilara et al. (2018). "The modulatory role of pre-SMA in speed-accuracy tradeoff: A bi-directional TMS study". In: *Neuropsychologia*. ISSN: 18733514. DOI: 10.1016/j.neuropsychologia.2017.12.031.

Berridge, Kent C (July 2009). "'Liking' and 'wanting' food rewards: Brain substrates and roles in eating disorders". In: *Physiology and Behavior* 97.5, pp. 537–550. ISSN: 00319384. DOI: 10.1016/j.physbeh.2009.02.044. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2717031&tool=pmcentrez&rendertype=abstract.

— (2004). "Motivation concepts in behavioral neuroscience". In: *Physiology and Behavior* 81.2, pp. 179–209. ISSN: 00319384. DOI: 10.1016/j.physbeh.2004.02.004.

Berridge, Kent C and Terry E Robinson (2003). *Parsing reward*. Vol. 26. 9. 507–513. ISBN: 0166-2236 (Print)$backslash$n0166-2236 (Linking). DOI: 10.1016/S0166-2236(03)00233-9.

Berridge, Kent C et al. (Feb. 2009). *Dissecting components of reward: 'liking', 'wanting', and learning*. Vol. 9. 1. 65–73. ISBN: 1471-4892 (Print)$backslash$n1471-4892 (Linking). DOI: 10.1016/j.coph.2008.12.014. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2756052&tool=pmcentrez&rendertype=abstract.

Bertsekas, Dimitri P et al. (1995). *Dynamic programming and optimal control*. Vol. 1. 2. Athena Scientific Belmont, MA.

Blanchard, Tommy C et al. (2015). "Orbitofrontal Cortex Uses Distinct Codes for Different Choice Attributes in Decisions Motivated by Curiosity Article Orbitofrontal Cortex Uses Distinct Codes for Different Choice Attributes in Decisions Motivated by Curiosity". In: *NEURON* 85.3, pp. 602–614. ISSN: 0896-6273. DOI: 10.1016/j.neuron.2014.12.050. URL: http://dx.doi.org/10.1016/j.neuron.2014.12.050.

Blythe, Mark a (2003). *Funology: from usability to enjoyment*. Vol. 3. 293 pp. ISBN: 1-4020-1252-7. DOI: 10.1007/1-4020-2967-5. URL: http://books.google.com/books?id=7Exv8MxJegMC&pgis=1.

Bogacz, Rafal and Rafal Bogacz (2007). "Optimal decision-making theories: linking neurobiology with behaviour". In: *Trends in Cognitive Sciences*. ISSN: 1364-6613. DOI: 10.1016/j.tics.2006.12.006.

Bogacz, Rafal et al. (2006). "The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks". In: *Psychological Review*. ISSN: 0033295X. DOI: 10.1037/0033-295X.113.4.700.

Bogacz, Rafal et al. (2010). "The neural basis of the speed–accuracy tradeoff". In: *Trends in Neurosciences*. ISSN: 01662236. DOI: 10.1016/j.tins.2009.09.002.

Böhmer, Matthias et al. (2011). "Falling asleep with Angry Birds, Facebook and Kindle". In: *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services - MobileHCI '11*. ISBN: 978-1-4503-0541-9. DOI: 10.1145/2037373.2037383.

Böhmer, Matthias et al. (2013a). "Appfunnel: A framework for usage-centric evaluation of recommender systems that suggest mobile applications". In: *Proceedings of the 2013 international conference on Intelligent user interfaces*. ACM, pp. 267–276.

Böhmer, Matthias et al. (2013b). "What's in the Apps for Context?: Extending a Sensor for Studying App Usage to Informing Context-awareness". In: *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*. UbiComp '13 Adjunct. New York, NY, USA: ACM, pp. 1423–1426. ISBN: 978-1-4503-2215-7. DOI: 10.1145/2494091.2496038. URL: http://doi.acm.org/10.1145/2494091.2496038.

Bolles, Robert (1975). *Theory of motivation.* HarperCollins Publishers.

Botvinick, Matthew M (2008). "Hierarchical models of behavior and prefrontal function". In: *Trends in cognitive sciences* 12.5, pp. 201–208.

Botvinick, Matthew M. et al. (2004). *Conflict monitoring and anterior cingulate cortex: An update*. ISBN: 1364-6613 (Print)$backslash$r1364-6613 (Linking). DOI: 10.1016/j.tics.2004.10.003.

Botvinick, Matthew M et al. (2009). "Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective". In: *Cognition* 113.3, pp. 262–280.

Botvinick, Matthew and Todd Braver (2015). "Motivation and cognitive control: from behavior to neural mechanism". In: *Annual Review of Psychology* 66, pp. 83–113.

Bouvier, Patrice et al. (2014). "Defining Engagement and Characterizing Engaged-Behaviors in Digital Gaming". In: *Simulation & Gaming* 45.4, pp. 491–507. DOI: 10.1177/1046878114553571.

Boyle, Elizabeth A et al. (2012). "Engagement in digital entertainment games: A systematic review". In: *Computers in human behavior* 28.3, pp. 771–780.

Boynton, G.M. (n.d.). *Washington Matlab Course*.

Bradley, Maragaret M et al. (2014). "Emotion and Motivation in the Perceptual Processing of Natural Scenes". In: *Scene Vision: Making Sense of What We See*, pp. 273–290. ISSN: 0262027852.

Brainard, David H. (1997). "The Psychophysics Toolbox". In: *Spatial Vision* 10.4, pp. 433–436. ISSN: 01691015. DOI: 10.1163/156856897X00357.

Braver, Todd S et al. (2014). "Mechanisms of motivation–cognition interaction: challenges and opportunities". In: *Cognitive, Affective, & Behavioral Neuroscience* 14.2, pp. 443–472.

Braziunas, Darius (2003). "Pomdp solution methods". In: *University of Toronto*.

Brosh, Allie (2013). *Hyperbole and a half: Unfortunate situations, flawed coping mechanisms, mayhem, and other things that happened*. Random House. ISBN: 978-1-4767-6459-7.

Brown, Alan S (1991). "A review of the tip-of-the-tongue experience". In: *Psychological bulletin* 109.2, pp. 204–223. ISSN: 00332909.

Busemeyer, J R and A Rapoport (1988). "Psychological models of deferred decision-making". In: *Journal of Mathematical Psychology* 32.2, pp. 91–134.

Carrasco, Marisa (2011). "Visual attention: The past 25 years". In: *Vision research* 51.13, pp. 1484–1525.

Carrier, L Mark et al. (2015). "Causes, effects, and practicalities of everyday multitasking". In: *Developmental Review* 35, pp. 64–78.

Carter, Evan C et al. (2015). "Reassessing intertemporal choice: human decision-making is more optimal in a foraging task than in a self-control task". In: *Frontiers in psychology* 6.

Carver, C S and M F Scheier (1981). *Attention and self-regulation: A control-theory approach to human behavior*. 1–403. ISBN: 0-387-90553-7. DOI: 10.1007/978-1-4612-5887-2.

— (1998). *On the self regulation of behavior*. 9. 1689–1699. ISBN: 978-85-7811-079-6.

Carver, Charles (2003). "Pleasure as a sign you can attend to something else: Placing positive feelings within a general model of affect". In: *Cognition & Emotion* 17.2, pp. 241–261.

Carver, Charles S. and Michael F. Scheier (Nov. 2002). "Control Processes and Self-Organization as Complementary Principles Underlying Behavior". In: *Personality and Social Psychology Review* 6.4, pp. 304–315. ISSN: 1088-8683. DOI: 10.1207/S15327957PSPR0604_05. URL: http://psr.sagepub.com/cgi/doi/10.1207/S15327957PSPR0604_05.

Cavanagh, James F. et al. (2011). "Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold". In: *Nature Neuroscience*. ISSN: 10976256. DOI: 10.1038/nn.2925.

Cavanagh, James F. et al. (2014). "Eye tracking and pupillometry are indicators of dissociable latent decision processes". In: *Journal of Experimental Psychology: General*. ISSN: 00963445. DOI: 10.1037/a0035813.

Charles, Lucie et al. (2013). "Distinct brain mechanisms for conscious versus subliminal error detection". In: *NeuroImage* 73, pp. 80–94. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2013.01.054. URL: http://dx.doi.org/10.1016/j.neuroimage.2013.01.054.

Charnov, Eric L (1976). "Optimal foraging, the marginal value theorem". In: *Theoretical population biology* 9.2, pp. 129–136.

Chater, N and M Oaksford (Feb. 1999). "Ten years of the rational analysis of cognition." In: *Trends in cognitive sciences* 3.2, pp. 57–65. ISSN: 1879-307X. URL: http://www.ncbi.nlm.nih.gov/pubmed/10234228.

Chiappori, Pierre-André and Arthur Lewbel (2015). "Gary Becker's a theory of the allocation of time". In: *The Economic Journal* 125.583, pp. 410–442.

Choe, Kyoung Whan et al. (2016). "Pupil size dynamics during fixation impact the accuracy and precision of video-based gaze estimation". In: *Vision Research* 118, pp. 48–59. ISSN: 18785646. DOI: 10.1016/j.visres.2014.12.018. URL: http://dx.doi.org/10.1016/j.visres.2014.12.018.

Christie, S. Thomas and Paul Schrater (2015). "Cognitive cost as dynamic allocation of energetic resources". In: *Frontiers in Neuroscience* 9, p. 289. ISSN: 1662-453X. DOI: 10.3389/fnins.2015.00289.

Cisek, P. et al. (2009). "Decisions in Changing Conditions: The Urgency-Gating Model". In: *Journal of Neuroscience*. ISSN: 00140015. DOI: 10.3112/erdkunde.2009.04.01.

Cisek, Paul (1999). "Beyond the computer metaphor: Behaviour as interaction". In: *Journal of Consciousness Studies* 6.11-12, pp. 125–142.

Clark, C W and M Mangel (2000). *Dynamic State Variable Models in Ecology*. 289 pp. ISBN: 0-19-512266-6. URL: http://books.google.com/books?id=CLi_wUxONE8C.

Cohen, Jonathan D et al. (May 2007). "Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration." In: *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 362.1481, pp. 933–42. ISSN: 0962-8436. DOI: 10.1098/rstb.2007.2098. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2430007&tool=pmcentrez&rendertype=abstract.

Colgan, P (1989). *Chapman and Hall animal behaviour series. Animal motivation.*

Collier, George H and Carolyn K Rovee-Collier (1983). "An ecological perspective of reinforcement and motivation". In: *Motivation*. Springer, pp. 427–441.

Collins, Anne and Etienne Koechlin (Jan. 2012). "Reasoning, learning, and creativity: Frontal lobe function and human decision-making". In: *PLoS Biology* 10.3, e1001293. ISSN: 15449173. DOI: 10.1371/journal.pbio.1001293. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3313946&tool=pmcentrez&rendertype=abstract.

Cooper, Richard P and Tim Shallice (2006). "Hierarchical schemas and goals in the control of sequential behavior." In: *Psychological review* 113.4, 887–916, discussion 917–931. ISSN: 0033-295X. DOI: 10.1037/0033-295X.113.4.887.

Corbetta, Maurizio and Gordon L. Shulman (2002). "Control of goal-directed and stimulus-driven attention in the brain". In: *Nature Reviews Neuroscience*. ISSN: 14710048. DOI: 10.1038/nrn755.

Cornwell, Benjamin et al. (2019). "The Social Structure of Time: Emerging Trends and New Directions". In: *Annual Review of Sociology* 45.

Cowan, Nelson (2010). "The magical mystery four: How is working memory capacity limited, and why?" In: *Current Directions in Psychological Science* 19.1, pp. 51–57. ISSN: 09637214. DOI: 10.1177/0963721409359277.

Cowley, Ben et al. (2008). "Toward an understanding of flow in video games". In: *Computers in Entertainment* 6.2, p. 1. ISSN: 15443574. DOI: 10.1145/1371216.1371223.

Cox, M et al. (2011). "Toward an Integrated Metacognitive Infrastructure". In: *2011 AAAI Fall Symposium*, pp. 74–81.

Csikszentmihalyi, Mihaly (1990). *Flow: The psychology of optimal performance*. ISBN: 0-521-34288-0.

Csikszentmihalyi, Mihaly and Judith Lefevre (1989). "Optimal Experience in Work and Leisure". In: *Journal of Personality and Social Psychology* 56.5, pp. 815–822. ISSN: 0022-3514. DOI: 10.1037/0022-3514.56.5.815.

D'Mello, Sidney and Art Graesser (2009). "Automatic detection of learner's affect from gross body language". In: *Applied Artificial Intelligence* 23.2, pp. 123–150.

— (2012). "Dynamics of affective states during complex learning". In: *Learning and Instruction* 22.2, pp. 145–157. ISSN: 09594752. DOI: 10.1016/j.learninstruc.2011.10.001.

D'Mello, Sidney et al. (2017). "Advanced, analytic, automated (AAA) measurement of engagement during learning". In: *Educational psychologist* 52.2, pp. 104–123.

Dabbish, Laura et al. (2011). “Why do i keep interrupting myself?: environment, habit and self-interruption”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, pp. 3127–3130.

Danziger, Kurt (1997). *Naming the mind: How psychology found its language*. Sage.

Daw, Nathaniel D et al. (Dec. 2005). “Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control.” In: *Nature neuroscience* 8.12, pp. 1704–11. ISSN: 1097-6256. DOI: 10.1038/nn1560. URL: http://www.ncbi.nlm.nih.gov/pubmed/16286932.

Daw, Nathaniel D et al. (June 2006). “Cortical substrates for exploratory decisions in humans”. In: *Nature* 441.7095, pp. 876–879. ISSN: 0028-0836. DOI: 10.1038/nature04766. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2635947&tool=pmcentrez&rendertype=abstract%20http://www.nature.com/doifinder/10.1038/nature04766.

Dawkins, R (1976). “Hierarchical organisation: A candidate principle for ethology”. In: *Growing points in ethology*, pp. 7–54.

Dayan, Peter and L. F. Abbott (2001). *Theoretical Neuroscience*. Cambridge, MA: MIT Press. 806 pp.

Dayan, Peter and Kent C Berridge (June 2014). “Model-based and model-free Pavlovian reward learning: Revaluation, revision, and revelation.” In: *Cognitive, affective & behavioral neuroscience* 14.2, pp. 473–92. ISSN: 1531-135X. DOI: 10.3758/s13415-014-0277-8.

De Vico Fallani, Fabrizio et al. (2010). “Defecting or not defecting: How to ”read” human behavior during cooperative games by EEG measurements”. In: *PLoS ONE* 5.12. ISSN: 19326203. DOI: 10.1371/journal.pone.0014187.

Demski, Andrés and Mariano Llamedo Soria (2016). “ecg-kit: a Matlab Toolbox for Cardiovascular Signal Processing . Journal of Open Research Software”. In: *Journal of Open Research Software* 4.1, e8. DOI: http://doi.org/10.5334/jors.86. URL: https://openresearchsoftware.metajnl.com/articles/10.5334/jors.86/.

Denison, Rachel N. et al. (2018). “Humans incorporate attention-dependent uncertainty into perceptual decisions and confidence”. In: *Proceedings of the National Academy of Sciences*, p. 201717720. ISSN: 0027-8424. DOI: 10.1073/pnas.1717720115. URL: http://www.pnas.org/lookup/doi/10.1073/pnas.1717720115.

Desai, Nishant et al. (2018). “Negotiable reinforcement learning for pareto optimal sequential decision-making”. In: *Advances in Neural Information Processing Systems*, pp. 4712–4720.

Dickey, Michele D and Sid Meier (2005). “Engaging By Design : How Engagement Strategies in Popular Computer and Video Games Can Inform Instructional Design”. In: 53.2, pp. 67–83.

Dickinson, A and B Balleine (2002). “The role of learning in the operation of motivational systems.” In: *Learning, Motivation and Emotion (Gallistel, C.R., ed.)* 3.

Dickinson, A and B Balleine (n.d.). “The role of learning in the operation of motivational systems.” In: *Learning, Motivation and Emotion (Gallistel, C.R., ed.)* 3 ().

Doll, Bradley B. et al. (2012). *The ubiquity of model-based reinforcement learning*. ISBN: 0959-4388. DOI: 10.1016/j.conb.2012.08.003.

Draheim, Christopher et al. (2016). “Combining reaction time and accuracy: The relationship between working memory capacity and task switching as a case example”. In: *Perspectives on Psychological Science* 11.1, pp. 133–155.

Drugowitsch, J. et al. (2012). “The Cost of Accumulating Evidence in Perceptual Decision Making”. In: *Journal of Neuroscience*. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.4010-11.2012.

Drugowitsch, Jan et al. (2014). “Relation between belief and performance in perceptual decision making”. In: *PLoS ONE* 9.5. ISSN: 19326203. DOI: 10.1371/journal.pone.0096511.

Ducheneaut, N. (Oct. 2006). “Building an MMO With Mass Appeal: A Look at Gameplay in World of Warcraft”. In: *Games and Culture* 1.4, pp. 281–317. ISSN: 1555-4120. DOI: 10.1177/1555412006292613. URL: http://gac.sagepub.com/cgi/doi/10.1177/1555412006292613.

Edge, Robert (2013). *Predicting Player Churn in Multiplayer Games using Goal-Weighted Empowerment*. Citeseer.

Einstein, Gilles O and Mark A McDaniel (2005). “Prospective memory: Multiple retrieval processes”. In: *Current Directions in Psychological Science* 14.6, pp. 286–290.

Ekirch, a R (2001). “Sleep we have lost: pre-industrial slumber in the British Isles.” In: *The American historical review*. ISSN: 0002-8762. DOI: 10.1086/ahr/106.2.343.

Ekman, Paul (1993). “Facial Expression and Emotion”. In: *American Psychologist*.

Eldar, Eran et al. (2016). “Mood as Representation of Momentum”. In: *Trends in Cognitive Sciences* 20.1, pp. 15–24. ISSN: 1879307X. DOI: 10.1016/j.tics.2015.07.010. URL: http://dx.doi.org/10.1016/j.tics.2015.07.010.

Elliot, Andrew J (2006). “The hierarchical model of approach-avoidance motivation”. In: *Motivation and Emotion* 30.2, pp. 111–116. ISSN: 01467239. DOI: 10.1007/s11031-006-9028-7.

Elliot, Andrew J. and James W. Fryer (2008). “The Goal Construct in Psychology”. In: *Handbook of motivation science*, pp. 235–247.

Escobar, Carolina et al. (2011). “Scheduled meals and scheduled palatable snacks synchronize circadian rhythms: Consequences for ingestive behavior”. In: *Physiology and Behavior* 104.4, pp. 555–561. ISSN: 00319384. DOI: 10.1016/j.physbeh.2011.05.001. URL: http://dx.doi.org/10.1016/j.physbeh.2011.05.001.

Fawcett, Tim W. et al. (Mar. 2012). “When is it adaptive to be patient? A general framework for evaluating delayed rewards”. In: *Behavioural Processes* 89.2, pp. 128–136. ISSN: 03766357. DOI: 10.1016/j.beproc.2011.08.015. URL: http://www.ncbi.nlm.

nih.gov/pubmed/21920413%20http://www.sciencedirect.com/science/article/pii/S0376635711001689.

Fawcett, Tim W et al. (Mar. 2014). "The evolution of decision rules in complex environments." In: *Trends in cognitive sciences* 18.3, pp. 153–161. ISSN: 1879-307X. DOI: 10.1016/j.tics.2013.12.012. URL: http://www.ncbi.nlm.nih.gov/pubmed/24467913.

Feldman, J and I Barshi (2007). *The Effects of Blood Glucose Levels on Cognitive Performance: A Review of Literature*. ISBN: 2-00-700317-1. URL: https://ntrs.nasa.gov/search.jsp?R=20070031714.

FeldmanHall, Oriel et al. (2016). "Emotion and decision-making under uncertainty: Physiological arousal predicts increased gambling during ambiguity but not risk." In: *Journal of Experimental Psychology: General* 145.10, pp. 1255–1262. ISSN: 1939-2222. DOI: 10.1037/xge0000205. URL: http://doi.apa.org/getdoi.cfm?doi=10.1037/xge0000205.

Feller, W (1968). *An Introduction to Probability Theory and Its Applications VOLUME I*. ISBN: 0-471-25708-7. DOI: 10.2307/1266435.

Fentress, John C (1983). "Ethological models of hierarchy and patterning of species-specific behavior". In: *Motivation*. Springer US, pp. 185–234.

Ferguson, Ts and Alex Cox (2012). "Optimal stopping and applications". In: *Optimal Stopping and Applications*. 2, pp. 1–32. URL: http://www.citeulike.org/group/4221/article/10559549.

Fetsch, Christopher R. et al. (2014). "Effects of Cortical Microstimulation on Confidence in a Perceptual Decision". In: *Neuron*. ISSN: 10974199. DOI: 10.1016/j.neuron.2014.07.011.

Fleming, S. M. et al. (2012). "Prefrontal Contributions to Metacognition in Perceptual Decision Making". In: *Journal of Neuroscience* 32.18, pp. 6117–6125. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.6489-11.2012. URL: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.6489-11.2012.

Fleming, Stephen M. et al. (2010). "Relating introspective accuracy to individual differences in brain structure". In: *Science* 329.5998, pp. 1541–1543. ISSN: 10959203. DOI: 10.1126/science.1191883.

Flood, Sarah M et al. (2018). "Daily temporal pathways: A Latent class approach to time diary data". In: *Social indicators research* 135.1, pp. 117–142.

Flower, Linda and John R. Hayes (1981). "A Cognitive Process Theory of Writing". In: *College Composition and Communication*. ISSN: 0010096X. DOI: 10.2307/356600.

Frank, Michael J. et al. (Jan. 2015). "fMRI and EEG Predictors of Dynamic Decision Parameters during Human Reinforcement Learning". In: *Journal of Neuroscience* 35.2, pp. 485–494. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.2036-14.2015. URL: http://www.jneurosci.org/content/35/2/485%20http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.2036-14.2015.

Gallistel, C R (1981). “Precis of Gallistel’s The organization of action: A new synthesis”. In: *The Behavioral and Brain Sciences* 4, pp. 609–650.
— (1990). *The organization of learning*. Vol. 3. 1112–1136. ISBN: 0-262-07113-4.
— (1978). “The irrelevance of past pleasure.” In: *Behavioral and Brain Sciences*.
Gamble, Katherine R. et al. (2018). “Different profiles of decision making and physiology under varying levels of stress in trained military personnel”. In: *International Journal of Psychophysiology*. ISSN: 18727697. DOI: 10.1016/j.ijpsycho.2018.03.017.
Gazzaley, Adam and Larry D Rosen (2016). *The Distracted Mind: Ancient Brains in a High-Tech World*. MIT Press.
Gerstel, Naomi and Dan Clawson (2018). “Control over time: Employers, workers, and families shaping work schedules”. In: *Annual review of sociology* 44, pp. 77–97.
Gittins, John et al. (2011). *Multi-Armed Bandit Allocation Indices: 2nd Edition*. ISBN: 978-0-470-67002-6. DOI: 10.1002/9780470980033.
Glennie, Paul and Nigel Thrift (1996). “Reworking EP Thompson’sTime, work-discipline and industrial capitalism’”. In: *Time & Society* 5.3, pp. 275–299.
Gollwitzer, Peter M. (1990). “Action phases and mind-sets”. In: *Handbook of motivation and cognition: Foundations of social behavior*, pp. 53–92.
Goodfellow, Ian et al. (2016). “Deep Learning”. In: *MIT Press*. ISSN: 1793-351X. DOI: 10.1142/S1793351X16500045.
Gopen, George D and Judith A Swan (1990). “The science of scientific writing”. In: *American Scientist*. ISSN: 0003-0996. DOI: 10.1007/978-0-387-77415-2.
Graus, David et al. (2016). “Analyzing and predicting task reminders”. In: *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*. ACM, pp. 7–15.
Green, D. G. and J. A. Swets (1966). *Signal detection theory and psychophysics*. ISBN: 0-932146-23-6. DOI: 10.1901/jeab.1969.12-475.
Greenberg, L M and I D Waldman (1993). “Developmental normative data on the test of variables of attention (T.O.V.A.).” In: *Journal of child psychology and psychiatry, and allied disciplines* 34, pp. 1019–1030. ISSN: 0021-9630.
Gross, Daniel R (1984). “Time allocation: A tool for the study of cultural behavior”. In: *Annual review of anthropology* 13.1, pp. 519–558.
Gross, James J and Lisa Feldman Barrett (2011). “Emotion Generation and Emotion Regulation: One or Two Depends on Your Point of View.” In: *Emotion review : journal of the International Society for Research on Emotion* 3.1, pp. 8–16. ISSN: 1754-0739. DOI: 10.1177/1754073910380974. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3072688&tool=pmcentrez&rendertype=abstract.
Hagger, Martin S. et al. (2016). “A Multilab Preregistered Replication of the Ego-Depletion Effect”. In: *Perspectives on Psychological Science*. ISSN: 17456924. DOI: 10.1177/1745691616652873.

Hamermesh, Daniel S et al. (2005). “Data watch: The American time use survey”. In: *Journal of Economic Perspectives* 19.1, pp. 221–232.

Harris, Tristan (2017). *Time Well Spent*. URL: http://www.timewellspent.io/.

Hartonen, T and M J Alava (2013). “How important tasks are performed : peer review”. In: *Scientific reports* 3, pp. 1–11. ISSN: 2045-2322. DOI: 10.1038/srep01679.

Hastie, Reid and Robyn M Dawes (2010). *Rational Choice in an Uncertain World: The Psychology of Judgement and Decision Making*. 2. 392 pp. ISBN: 0-7619-2275-X.

Hauser, Tobias U et al. (2016). “Computational psychiatry of ADHD: neural gain impairments across Marrian levels of analysis”. In: *Trends in neurosciences* 39.2, pp. 63–73.

Hayden, Benjamin Y et al. (July 2011). “Neuronal basis of sequential foraging decisions in a patchy environment.” In: *Nature neuroscience* 14.7, pp. 933–9. ISSN: 1546-1726. DOI: 10.1038/nn.2856. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3553855&tool=pmcentrez&rendertype=abstract.

Heckman, James J (2015). “Introduction to a Theory of the Allocation of Time by Gary Becker”. In: *The Economic Journal* 125.583, pp. 403–409.

Herbert, Simon et al. (1962). “The architecture of complexity”. In: *Proceedings of the American Philosophical Society* 106.6, pp. 467–482.

Hobbs, Malcolm et al. (Apr. 2005). “Dissociation of wanting and liking for alcohol in humans: A test of the incentive-sensitisation theory”. In: *Psychopharmacology* 178.4, pp. 493–499. ISSN: 00333158. DOI: 10.1007/s00213-004-2026-0. URL: http://www.ncbi.nlm.nih.gov/pubmed/15517194.

Hoffman, Bobby and Louis Nadelson (2010). “Motivational engagement and video gaming: A mixed methods study”. In: *Educational Technology Research and Development* 58.3, pp. 245–270.

Hofmann, Wilhelm et al. (2012). “Executive functions and self-regulation”. In: *Trends in Cognitive Sciences* 16.3, pp. 174–180. ISSN: 13646613. DOI: 10.1016/j.tics.2012.01.006. URL: http://dx.doi.org/10.1016/j.tics.2012.01.006.

Houston, Alasdair I and John M McNamara (Jan. 2014). “Foraging currencies, metabolism and behavioural routines.” In: *The Journal of animal ecology* 83.1, pp. 30–40. ISSN: 1365-2656. DOI: 10.1111/1365-2656.12096. URL: http://www.ncbi.nlm.nih.gov/pubmed/23730810.

Houston, A et al. (1988). “dynamic models in behavioural and evolutionary ecology”. In: *Nature*. URL: http://people.ucsc.edu/%20msmangel/Nature%201988.pdf.

Huizenga, J et al. (2010). “Claims about games: A literature review of a decade of research on the effects on learning and motivation”. In: *Proceedings of the 4th European Conference on Games Based Learning. Reading, UK: Academic Conferences Ltd.*

Hull, Clark L. (1943). “Principles of Behavior: An Introduction to Behavior Theory.” In: *The Journal of Abnormal and Social Psychology*. Vol. 39. 3, pp. 377–380. DOI: 10.1037/h0051597. URL: http://doi.apa.org/getdoi.cfm?doi=10.1037/h0051597.

Iqbal, Shamsi T and Brian P Bailey (2008). "Effects of intelligent notification management on users and their tasks". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, pp. 93–102.

Isaksen, Aaron et al. (2015). "Exploring Game Space Using Survival Analysis." In: *FDG*.

Jagacinski, Richard J and John M Flach (2003). *Control theory for humans: Quantitative approaches to modeling performance*. CRC Press.

Jara-Díaz, Sergio and Jorge Rosales-Salas (2017). "Beyond transport time: A review of time use modeling". In: *Transportation Research Part A: Policy and Practice* 97, pp. 209–230.

Jarmolowicz, David P et al. (2010). "Temporal patterns of behavior from the scheduling of psychology quizzes". In: *Journal of applied behavior analysis* 43.2, pp. 297–301.

Jin, Jing and Laura A Dabbish (2009). "Self-interruption on the computer: a typology of discretionary task interleaving". In: *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, pp. 1799–1808.

Jo, Hang Hyun et al. (2012). "Time-varying priority queuing models for human dynamics". In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 85.6, pp. 1–8. ISSN: 15393755. DOI: 10.1103/PhysRevE.85.066101.

Jong, Sijmen de (1993). "SIMPLS: An alternative approach to partial least squares regression". In: *Chemometrics and Intelligent Laboratory Systems*. ISSN: 01697439. DOI: 10.1016/0169-7439(93)85002-X.

Jönsson, Fredrik U. et al. (2005). "Odor emotionality affects the confidence in odor naming". In: *Chemical Senses* 30.1, pp. 29–35. ISSN: 0379864X. DOI: 10.1093/chemse/bjh254.

Judd, Terry (2015). "Task selection, task switching and multitasking during computer-based independent study". In: *Australasian Journal of Educational Technology* 31.2.

Jung, Tobias et al. (2012). "Empowerment for Continuous Agent-Environment Systems". In: *Adaptive Behavior* 19, pp. 16–39. URL: http://arxiv.org/abs/1201.6583.

Juslin, Peter et al. (2000). "Naive empiricism and dogmatism in confidence research: A critical examination of the hard-easy effect". In: *Psychological Review* 107.2, pp. 384–396. ISSN: 0033295X. DOI: 10.1037/0033-295X.107.2.384.

Kahneman, Daniel and Alan B Krueger (2006a). "American Economic Association Developments in the Measurement of Subjective Well-Being". In: *Source: The Journal of Economic Perspectives* 20.1, pp. 3–24. URL: http://www.jstor.org/stable/30033631%5Cnhttp://about.jstor.org/terms.

— (2006b). "Developments in the Measurement of Subjective Well-Being". In: *Journal of Economic Literature* 20.1, pp. 3–24.

Kane, Michael J et al. (2007). "For Whom the Mind Wanders, and When". In: *Psychological Science* 18.7, pp. 614–621. ISSN: 0956-7976. DOI: 10.1111/j.1467-9280.2007.01948.x.

Kapoor, Komal (2014). "Models of Dynamic User Preferences and their Applications to Recommendation and Retention". PhD thesis.

Kapoor, Komal et al. (2015). "Just in time recommendations: Modeling the dynamics of boredom in activity streams". In: *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*. ACM, pp. 233–242.

Karonen, Ilmari (2014). *Response to Estimating rate of occurrence of an event with exponential smoothing and irregular events*. `https://stackoverflow.com/a/23617678`. Accessed: 2019-11-12.

Kepecs, Adam and Zachary F. Mainen (2014). "A computational framework for the study of confidence across species". In: *The Cognitive Neuroscience of Metacognition* 9783642451, pp. 115–145. ISSN: 14712970. DOI: `10.1007/978-3-642-45190-4_6`.

Kepecs, Adam et al. (2008). "Neural correlates, computation and behavioural impact of decision confidence". In: *Nature* 455.7210, pp. 227–231. ISSN: 00280836. DOI: `10.1038/nature07200`.

Keramati, Mehdi and Boris Gutkin (2014). "Homeostatic reinforcement learning for integrating reward collection and physiological stability". In: *Elife* 3, e04811. DOI: `10.7554/eLife.04811`.

Kersten, Daniel et al. (1996). "Illusory motion from shadows." In:

Kersten, Daniel et al. (2004). "Object perception as Bayesian inference". In: *Annu. Rev. Psychol.* 55, pp. 271–304.

Kessler, Yoav et al. (June 2009). "Choosing to switch: spontaneous task switching despite associated behavioral costs." In: *Acta psychologica* 131.2, pp. 120–8. ISSN: 1873-6297. DOI: `10.1016/j.actpsy.2009.03.005`. URL: `http://www.ncbi.nlm.nih.gov/pubmed/19386295`.

Khajah, Mohammad M et al. (2016). "Designing engaging games using bayesian optimization". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, pp. 5571–5582.

Kiani, Roozbeh and Michael N Shadlen (2009). "Representation of Confidence Associated with a Decision by Neurons in the Parietal Cortex". In: *Science* 324.5928, pp. 759–764. ISSN: 0036-8075, 1095-9203. DOI: `10.1126/science.1169405`. URL: `http://www.sciencemag.org/content/324/5928/759`.

Kiani, Roozbeh et al. (Dec. 2014). "Choice certainty is informed by both evidence and decision time." In: *Neuron* 84.6, pp. 1329–42. ISSN: 1097-4199. URL: `http://www.cell.com/article/S0896627314010964/fulltext`.

Kidd, Celeste et al. (2013). "Rational snacking: Young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability". In: *Cognition* 126.1, pp. 109–114.

Kilpatrick, Lisa and Larry Cahill (2003). "Amygdala modulation of parahippocampal and frontal regions during emotionally influenced memory storage". In: *NeuroImage*. ISSN: 10538119. DOI: `10.1016/j.neuroimage.2003.08.006`.

Kim, Jinhong et al. (2013). “Microscopic Modelling Circadian and Bursty Pattern of Human Activities”. In: *PLoS ONE* 8.3, pp. 1–7. ISSN: 19326203. DOI: 10.1371/journal.pone.0058292.

Kimmig, Angelika et al. (2012). “A Short Introduction to Probabilistic Soft Logic”. In: *Proceedings of the NIPS Workshop on Probabilistic Programming: Foundations and Applications*.

Klyubin, Alexander S et al. (2005). “Empowerment : A Universal Agent-Centric Measure of Control”. In: *Evolutionary Computation, 2005. The 2005 IEEE Congress on*, pp. 128–135. ISBN: 0-7803-9363-5.

Koch, Iring et al. (2018). “Cognitive structure, flexibility, and plasticity in human multitasking-an integrative review of dual-task and task-switching research”. In: *Psychological Bulletin*. ISSN: 00332909. DOI: 10.1037/bul0000144.

Koechlin, Etienne et al. (Nov. 2003). “The Architecture of Cognitive Control in the Human Prefrontal Cortex”. In: *Science* 302.5648, pp. 1181–1185. ISSN: 00368075. DOI: 10.1126/science.1088545. URL: http://www.ncbi.nlm.nih.gov/pubmed/14615530.

Koo, Minjung and Ayelet Fishbach (2012). “The Small-Area Hypothesis: Effects of Progress Monitoring on Goal Adherence”. In: *Journal of Consumer Research* 39.3, pp. 493–509. ISSN: 00935301. DOI: 10.1086/663827. URL: http://www.jstor.org/stable/10.1086/663827.

Kool, Wouter and Matthew Botvinick (2018). “Mental labour”. In: *Nature human behaviour*, p. 1.

Kool, Wouter et al. (Nov. 2010). “Decision making and the avoidance of cognitive demand.” In: *Journal of experimental psychology. General* 139.4, pp. 665–82. ISSN: 1939-2222. DOI: 10.1037/a0020198. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2970648&tool=pmcentrez&rendertype=abstract.

Körding, Konrad P and Daniel M Wolpert (2006). “Bayesian decision theory in sensorimotor control”. In: *Trends in cognitive sciences* 10.7, pp. 319–326.

Koriat, Asher (2012). “The self-consistency model of subjective confidence”. In: *Psychological Review* 119.1, pp. 80–113. ISSN: 0033295X. DOI: 10.1037/a0025648.

Koutstaal, Wilma (2012). *The agile mind*. Oxford University Press.

Krishnan, Anjali et al. (2011). “Partial Least Squares (PLS) methods for neuroimaging: A tutorial and review”. In: *NeuroImage*. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2010.07.034.

Kulkarni, Tejas D et al. (2016). “Hierarchical deep reinforcement learning: Integrating temporal abstraction and intrinsic motivation”. In: *Advances in neural information processing systems*, pp. 3675–3683.

Kunimoto, Craig et al. (2001). “Confidence and accuracy of near-threshold discrimination responses”. In: *Consciousness and Cognition* 10.3, pp. 294–340. ISSN: 10538100. DOI: 10.1006/ccog.2000.0494.

Kurby, Christopher A. and Jeffrey M. Zacks (2008). *Segmentation in the perception and memory of events*. Vol. 12. 2. 72–79. ISBN: 1364-6613. DOI: 10.1016/j.tics.2007.11.004.

Kurzban, Robert et al. (2013). "An opportunity cost model of subjective effort and task performance". In: *Behavioral and brain sciences* 36.6, pp. 661–679.

Laeng, Bruno et al. (2012). "Pupillometry: A window to the preconscious?" In: *Perspectives on Psychological Science*. ISSN: 17456916. DOI: 10.1177/1745691611427305.

Lake, Brenden M et al. (2017). "Building machines that learn and think like people". In: *Behavioral and brain sciences* 40.

Lange, Floris P. de et al. (2011). "How awareness changes the relative weights of evidence during human decision-making". In: *PLoS Biology* 9.11. ISSN: 15449173. DOI: 10.1371/journal.pbio.1001203.

Lee, John D et al. (2013). *The Oxford handbook of cognitive engineering*. Oxford University Press.

Lee, Yoonjoo et al. (2016). "Reliability, validity, and variability of the subjective well-being questions in the 2010 American Time Use Survey". In: *Social indicators research* 126.3, pp. 1355–1373.

Leite, Fabio and Roger Ratcliff (2010). "Modeling reaction time and accuracy of multiple-alternative decisions". In: *Attention, perception & psychophysics* 72.3, pp. 246–273. ISSN: 1943-3921. DOI: 10.3758/APP.

Leiva, Luis et al. (2012). "Back to the App: The Costs of Mobile Application Interruptions". In: *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services*. MobileHCI '12. New York, NY, USA: ACM, pp. 291–294. ISBN: 978-1-4503-1105-2. DOI: 10.1145/2371574.2371617. URL: http://doi.acm.org/10.1145/2371574.2371617.

Lewis, Chris (2014). *Irresistible Apps: Motivational design patterns for apps, games, and web-based communities*. Apress.

Lieder, Falk and Thomas L Griffiths (2019). "Resource-rational analysis: understanding human cognition as the optimal use of limited computational resources". In: *Behavioral and Brain Sciences*, pp. 1–85.

Lima, Steven L. and Lawrence M. Dill (1990). "Behavioral decisions made under the risk of predation: a review and prospectus". In: *Canadian Journal of Zoology* 68.4, pp. 619–640. ISSN: 0008-4301. DOI: 10.1139/z90-092. URL: http://www.nrcresearchpress.com/doi/abs/10.1139/z90-092.

Litman, Jordan A. (Sept. 2005). "Curiosity and the pleasures of learning: Wanting and liking new information". In: *Cognition & Emotion* 19.6, pp. 793–814. ISSN: 0269-9931. DOI: 10.1080/02699930541000101. URL: http://www.tandfonline.com/doi/abs/10.1080/02699930541000101.

Litt, Ab et al. (Jan. 2010). "Lusting while loathing: parallel counterdriving of wanting and liking." In: *Psychological science : a journal of the American Psychological Society /*

APS 21.1, pp. 118–125. ISSN: 0956-7976. DOI: 10.1177/0956797609355633. URL: http://www.ncbi.nlm.nih.gov/pubmed/20424032.

Lo, Andreas et al. (2008). “Both Predator and Prey: Emotional Arousal in Threat and Reward”. In: *Psychological Science*, pp. 865–873.

Löchtefeld, Markus et al. (2013). “AppDetox: Helping Users with Mobile App Addiction”. In: *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia*. MUM ’13. New York, NY, USA: ACM, 43:1–43:2. ISBN: 978-1-4503-2648-3. DOI: 10.1145/2541831.2541870. URL: http://doi.acm.org/10.1145/2541831.2541870.

Locke, Edwin A. and Gary P. Latham (2006). “New directions in goal-setting theory”. In: *Current Directions in Psychological Science* 15.5, pp. 265–268. ISSN: 09637214. DOI: 10.1111/j.1467-8721.2006.00449.x.

Loewenstein, George (1994). “The psychology of curiosity: A review and reinterpretation.” In: *Psychological bulletin* 116.1, pp. 75–98. ISSN: 0033-2909. DOI: 10.1037//0033-2909.116.1.75. URL: http://doi.apa.org/getdoi.cfm?doi=10.1037/0033-2909.116.1.75%20http://psycnet.apa.org/journals/bul/116/1/75/.

Luce, R. Duncan (1986). *Response Times: Their Role in Inferring Elementary Mental Organization*. ISBN: 978-0-19-986987-9. DOI: 10.1093/acprof:oso/9780195070019.001.0001.

Ma, W J et al. (2014). “Changing concepts of working memory”. In: *Nature Neuroscience*. ISSN: 1546-1726; 1097-6256. DOI: 10.1038/nn.3655.

Mamassian, Pascal (2016). “Visual Confidence”. In: *Annual Review of Vision Science*. ISSN: 2374-4642. DOI: 10.1146/annurev-vision-111815-114630.

Mangel, M and CW Clark (1986). “Towards a unified foraging theory”. In: *Ecology*. URL: http://www.jstor.org/stable/1938669.

Mani, Anandi et al. (2013). “Poverty impedes cognitive function”. In: *science* 341.6149, pp. 976–980.

Mann, Traci et al. (May 2013). “Self-regulation of health behavior: social psychological approaches to goal setting and goal striving.” In: *Health psychology : official journal of the Division of Health Psychology, American Psychological Association* 32.5, pp. 487–98. ISSN: 1930-7810. DOI: 10.1037/a0028533. URL: http://www.ncbi.nlm.nih.gov/pubmed/23646832.

Mark, Gloria et al. (2015). “Focused, aroused, but so distractible: Temporal perspectives on multitasking and communications”. In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, pp. 903–916.

Mark, Gloria et al. (2018). “Effects of Individual Differences in Blocking Workplace Distractions”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, p. 92.

Marr, David (1982). *Vision: A computational investigation into the human representation and processing of visual information*. The MIT Press.

Marulanda-Carter, Laura and Thomas W. Jackson (2012). "Effects of e-mail addiction and interruptions on employees". In: *Journal of Systems and Information Technology* 14.1, pp. 82–94. ISSN: 1328-7265. DOI: 10.1108/13287261211221146.

Mata, Jutta et al. (2009). "Motivational "spill-over" during weight control: Increased self-determination and exercise intrinsic motivation predict eating self-regulation." In:

Matthews, Gerald et al. (2010). "Task engagement, cerebral blood flow velocity, and diagnostic monitoring for sustained attention." In: *Journal of experimental psychology. Applied* 16.2, pp. 187–203. ISSN: 1076-898X. DOI: 10.1037/a0019572.

Matzke, Dora and Eric Jan Wagenmakers (2009). "Psychological interpretation of the ex-gaussian and shifted wald parameters: A diffusion model analysis". In: *Psychonomic Bulletin and Review*. ISSN: 10699384. DOI: 10.3758/PBR.16.5.798.

McClure, Samuel M et al. (2003). "Temporal prediction errors in a passive learning task activate human striatum". In: *Neuron* 38.2, pp. 339–346.

McDaniel, R. and J. Fanfarelli (2016). "Building Better Digital Badges: Pairing Completion Logic With Psychological Factors". In: *Simulation & Gaming* 47.1, pp. 73–102. ISSN: 1046-8781. DOI: 10.1177/1046878115627138.

McIntosh, Anthony Randal and Nancy J. Lobaugh (2004). "Partial least squares analysis of neuroimaging data: Applications and advances". In: *NeuroImage*. ISBN: 1053-8119. DOI: 10.1016/j.neuroimage.2004.07.020.

Mehlhorn, Katja et al. (2015). "Unpacking the exploration–exploitation tradeoff: A synthesis of human and animal literatures." In: *Decision* 2.3, p. 191.

Melamed, S and S Bruhis (1996). "The effects of chronic industrial noise exposure on urinary cortisol, fatigue and irritability: a controlled field experiment." In: *Journal of occupational and environmental medicine / American College of Occupational and Environmental Medicine*. ISSN: 1076-2752. DOI: 10.1097/00043764-199603000-00009.

Merkle, Edgar C. (2009). "The disutility of the hard-easy effect in choice confidence". In: *Psychonomic Bulletin and Review* 16.1, pp. 204–213. ISSN: 10699384. DOI: 10.3758/PBR.16.1.204.

Meyniel, Florent et al. (2015). "Confidence as Bayesian Probability: From Neural Origins to Behavior". In: *Neuron* 88.1, pp. 78–92. ISSN: 10974199. DOI: 10.1016/j.neuron.2015.09.039. URL: http://dx.doi.org/10.1016/j.neuron.2015.09.039.

Miller, E K and J D Cohen (2001). "An integrative theory of prefrontal cortex function." In: *Annual review of neuroscience* 24, pp. 167–202. ISSN: 0147-006X. DOI: 10.1146/annurev.neuro.24.1.167.

Miller, George A et al. (1960). "Plans and the Structure of Behavior". In: *New York: Holt, Rinehart and Winston*.

Milosavljevic, Mili et al. (2010). "The Drift Diffusion Model Can Account for the Accuracy and Reaction Time of Value-Based Choices Under High and Low Time Pressure". In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.1901533.

Mischel, W et al. (1972). "Cognitive and attentional mechanisms in delay of gratification." In: *Journal of personality and social psychology* 21, pp. 204–218. ISSN: 0022-3514. DOI: 10.1037/h0032198.

Mischel, W et al. (1989). "Delay of gratification in children." In: *Science (New York, N.Y.)* 244, pp. 933–938. ISSN: 0036-8075. DOI: 10.1126/science.2658056.

Miyake, A et al. (2000). "The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis." In: *Cognitive psychology* 41.1, pp. 49–100. ISSN: 0010-0285. DOI: 10.1006/cogp.1999.0734.

Mobbs, Dean et al. (Mar. 2015). "The ecology of human fear: Survival optimization and the nervous system". In: *Frontiers in Neuroscience* 9 (FEB), pp. 1–22. ISSN: 1662453X. DOI: 10.3389/fnins.2015.00055. URL: http://www.frontiersin.org/Evolutionary_Psychology_and_Neuroscience/10.3389/fnins.2015.00055/abstract.

Mohamed, Shakir and Danilo Jimenez Rezende (2015). "Variational information maximisation for intrinsically motivated reinforcement learning". In: *Advances in neural information processing systems*, pp. 2125–2133.

Monsell, Stephen (2003). "Task switching". In: *Trends in cognitive sciences* 7.3, pp. 134–140.

Moore, Don A. and Paul J. Healy (2008). "The Trouble With Overconfidence". In: *Psychological Review* 115.2, pp. 502–517. ISSN: 0033295X. DOI: 10.1037/0033-295X.115.2.502.

Moskowitz, Gordon B and Heidi Grant (2009). *The psychology of goals*. Guilford press.

Murray, Greg et al. (2009). "Nature's clocks and human mood: the circadian system modulates reward motivation." In: *Emotion (Washington, D.C.)* 9.5, pp. 705–716. ISSN: 1528-3542. DOI: 10.1037/a0017080.

Nachum, Ofir et al. (2018). "Data-efficient hierarchical reinforcement learning". In: *Advances in Neural Information Processing Systems*, pp. 3303–3313.

Nassar, Matthew R. et al. (2012). "Rational regulation of learning dynamics by pupil-linked arousal systems". In: *Nature Neuroscience* 15.7, pp. 1040–1046. ISSN: 10976256. DOI: 10.1038/nn.3130.

Navajas, Joaquin et al. (2016). *Post-decisional accounts of biases in confidence*. DOI: 10.1016/j.cobeha.2016.05.005.

Neal, Andrew et al. (2017). "Dynamic self-regulation and multiple-goal pursuit". In: *Annual Review of Organizational Psychology and Organizational Behavior* 4, pp. 401–423.

Newell, Allen and Stuart K Card (1985). "The prospects for psychological science in human-computer interaction". In: *Human-computer interaction* 1.3, pp. 209–242.

Newsome, William T. et al. (Sept. 1989). "Neuronal correlates of a perceptual decision". In: *Nature* 341.6237, pp. 52–54. ISSN: 0028-0836. DOI: 10.1038/341052a0. URL: http://www.nature.com/doifinder/10.1038/341052a0.

Nieto-Barajas, Luis E. and Stephen G. Walker (2002). "Markov beta and gamma processes for modelling hazard rates". In: *Scandinavian Journal of Statistics*. ISSN: 03036898. DOI: 10.1111/1467-9469.00298.

Niv, Yael et al. (2006). "A normative perspective on motivation". In: *Trends in Cognitive Sciences* 10.8, pp. 375–381. ISSN: 13646613. DOI: 10.1016/j.tics.2006.06.010.

Niv, Yael et al. (2007). "Tonic dopamine: Opportunity costs and the control of response vigor". In: *Psychopharmacology* 191, pp. 507–520. ISSN: 00333158. DOI: 10.1007/s00213-006-0502-4.

Oaten, Allan (Dec. 1977). "Optimal foraging in patches: a case for stochasticity." In: *Theoretical Population Biology* 12.3, pp. 263–285. ISSN: 0040-5809. DOI: 10.1016/0040-5809(77)90046-6. URL: http://www.ncbi.nlm.nih.gov/pubmed/564087.

Oostenveld, Robert et al. (2011). "FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data". In: *Computational Intelligence and Neuroscience*. ISSN: 16875265. DOI: 10.1155/2011/156869.

Oudeyer, Pierre-Yves Yves and Frederic Kaplan (Jan. 2007). "What is Intrinsic Motivation? A Typology of Computational Approaches." In: *Frontiers in neurorobotics* 1 (November), p. 6. ISSN: 1662-5218. DOI: 10.3389/neuro.12.006.2007. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2533589&tool=pmcentrez&rendertype=abstract.

Ouelhadj, Djamila and Sanja Petrovic (2009). *A survey of dynamic scheduling in manufacturing systems*. Vol. 12. 4. 417–431. ISBN: 10946136. DOI: 10.1007/s10951-008-0090-8.

Palmer, John et al. (2005). "The effect of stimulus strength on the speed and accuracy of a perceptual decision". In: *Journal of Vision*. ISSN: 1534-7362. DOI: 10.1167/5.5.1.

Parate, Abhinav et al. (2013). "Practical prediction and prefetch for faster access to applications on mobile phones". In: *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM, pp. 275–284.

Paulus, Martin P and Angela J Yu (Sept. 2012). *Emotion and decision-making: Affect-driven belief systems in anxiety and depression*. Vol. 16. 9. 476 pp. ISBN: 1879-307X (Electronic)$backslash$n1364-6613 (Linking). DOI: 10.1016/j.tics.2012.07.009. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3446252&tool=pmcentrez&rendertype=abstract.

Payne, Stephen J et al. (2007). "Discretionary task interleaving: Heuristics for time allocation in cognitive foraging". In: *JOURNAL OF EXPERIMENTAL PSYCHOLOGY GENERAL* 136.3, p. 370.

Peirce, C and J Jastrow (1884). "On small differences of sensation". In: *Mem. Natl Acad. Sci.* ISSN: 0009-4978. DOI: 52,281-302.

Pentland, Wendy E et al. (1999). *Time use research in the social sciences*. Springer.

Persaud, Navindra et al. (2007). "Post-decision wagering objectively measures awareness". In: *Nature Neuroscience*. ISSN: 1520510X. DOI: 10.1021/acs.inorgchem.6b00522.

Peters, Megan A.K. et al. (2017). "Transcranial magnetic stimulation to visual cortex induces suboptimal introspection". In: *Cortex*. ISSN: 19738102. DOI: 10.1016/j.cortex.2017.05.017.

Phelps, Elizabeth A. (2006). "Emotion and Cognition: Insights from Studies of the Human Amygdala". In: *Annual Review of Psychology*. ISSN: 0066-4308. DOI: 10.1146/annurev.psych.56.091103.070234.

Phelps, Elizabeth A. and Joseph E. LeDoux (2005). *Contributions of the amygdala to emotion processing: From animal models to human behavior*. ISBN: 0896-6273 n0896-6273. DOI: 10.1016/j.neuron.2005.09.025.

Phelps, Elizabeth A. et al. (2014). "Emotion and Decision Making: Multiple Modulatory Neural Circuits". In: *Annual Review of Neuroscience* 37.1, pp. 263–287. ISSN: 0147-006X. DOI: 10.1146/annurev-neuro-071013-014119. URL: http://www.annualreviews.org/doi/10.1146/annurev-neuro-071013-014119.

Piech, Chris et al. (2015). "Deep knowledge tracing". In: *Advances in Neural Information Processing Systems*, pp. 505–513.

Pinedo, Michael (2012). *Scheduling*. Springer.

Pirolli, Peter L T (2007). "Information Foraging Theory". In: *World Wide Web Internet And Web Information Systems* 1, pp. 759–760. ISSN: 00071013. DOI: 10.1093/acprof:oso/9780195173321.001.0001. URL: http://www.oxfordscholarship.com/oso/public%20/content/psychology/9780195173321/toc.html.

Pirolli, Peter and Stuart Card (1995). "Information foraging in information access environments". In: *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '95*, pp. 51–58. DOI: 10.1145/223904.223911. URL: http://portal.acm.org/citation.cfm?doid=223904.223911.

Poldrack, Russell A. (2006). *Can cognitive processes be inferred from neuroimaging data?* ISBN: 1364-6613 (Print)$backslash$r1364-6613 (Linking). DOI: 10.1016/j.tics.2005.12.004.

Pouget, Alexandre et al. (2016). "Confidence and certainty: distinct probabilistic quantities for different goals". In: *Nature neuroscience* 11.10, pp. 1651–1666. ISSN: 1527-5418. DOI: 10.1080/10937404.2015.1051611.INHALATION.

Preuschoff, Kerstin et al. (2011). "Pupil dilation signals surprise: Evidence for noradrenaline's role in decision making". In: *Frontiers in Neuroscience* 5, pp. 1–12. ISSN: 16624548. DOI: 10.3389/fnins.2011.00115.

Proekt, Alex et al. (2012). "Scale invariance in the dynamics of spontaneous behavior." In: *Proceedings of the National Academy of Sciences of the United States of America* 109.26, pp. 10564–9. ISSN: 1091-6490. DOI: 10.1073/pnas.1206894109. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3387096&tool=pmcentrez&rendertype=abstract.

Rae, Babette et al. (2014). "The hare and the tortoise: Emphasizing speed can change the evidence used to make decisions". In: *Journal of Experimental Psychology: Learn-*

ing Memory and Cognition 40.5, pp. 1226–1243. ISSN: 02787393. DOI: 10.1037/a0036801.

Rafati, Jacob and David C Noelle (2019). "Learning representations in model-free hierarchical reinforcement learning". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 33, pp. 10009–10010.

Rahnev, D. A. et al. (2012). "Direct injection of noise to the visual cortex decreases accuracy but increases decision confidence". In: *Journal of Neurophysiology*. ISSN: 0022-3077. DOI: 10.1152/jn.00985.2011.

Rao, Vinayak and Yee Whye Teh (2013). "Fast MCMC sampling for Markov jump processes and extensions". In: *The Journal of Machine Learning Research* 14.1, pp. 3295–3320.

Ratcliff, R. and H. P. A. Van Dongen (2011). "Diffusion model for one-choice reaction-time tasks and the cognitive effects of sleep deprivation". In: *Proceedings of the National Academy of Sciences*. ISSN: 0027-8424. DOI: 10.1073/pnas.1100483108.

Ratcliff, Roger (1978). "A theory of memory retrieval." In: *Psychological Review* 85.2, pp. 59–108. ISSN: 0033-295X. DOI: 10.1037/0033-295X.85.2.59.

Ratcliff, Roger and Gail McKoon (2008). "The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks". In: *Neural Computation* 20.4, pp. 873–922. ISSN: 0899-7667. DOI: 10.1162/neco.2008.12-06-420. URL: http://www.mitpressjournals.org/doi/10.1162/neco.2008.12-06-420.

Ratcliff, Roger and Philip L. Smith (2004). "A Comparison of Sequential Sampling Models for Two-Choice Reaction Time". In: *Psychol Rev*, pp. 333–367. ISSN: 08966273. DOI: 10.1097/MCA.0000000000000178.Endothelial.

Ratcliff, Roger et al. (Apr. 2009). "Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG." In: 106.16, pp. 6539–44. ISSN: 1091-6490. DOI: 10.1073/pnas.0812589106.

Ratcliff, Roger et al. (2016). "Diffusion Decision Model: Current Issues and History". In: *Trends in Cognitive Sciences* 20.4, pp. 260–281. ISSN: 1879307X. DOI: 10.1016/j.tics.2016.01.007. URL: http://dx.doi.org/10.1016/j.tics.2016.01.007.

Reiss, Steven (2004). "Multifaceted nature of intrinsic motivation: The theory of 16 basic desires." In: *Review of general psychology* 8.3, p. 179.

Reschly, Amy L. and Sandra L. Christenson (2012). "Jingle, Jangle, and Conceptual Haziness: Evolution and Future Directions of the Engagement Construct". In: *Handbook of Research on Student Engagement*. Ed. by Sandra L. Christenson et al. Boston, MA: Springer US, pp. 3–19. ISBN: 978-1-4614-2018-7. DOI: 10.1007/978-1-4614-2018-7_1.

Robinson, Mike J F and Kent C Berridge (2013). "Instant transformation of learned repulsion into motivational "wanting"". In: *Current Biology* 23.4, pp. 282–289.

Roijers, Diederik M et al. (2013). "A survey of multi-objective sequential decision-making". In: *Journal of Artificial Intelligence Research* 48, pp. 67–113.

Rolls, Edmund T. (2009). *Emotion Explained*. 1–632. ISBN: 978-0-19-169379-3. DOI: 10.1093/acprof:oso/9780198570035.001.0001.

Rosen, Larry D. et al. (2013). "Facebook and texting made me do it: Media-induced task-switching while studying". In: *Computers in Human Behavior* 29.3, pp. 948–958. ISSN: 07475632. DOI: 10.1016/j.chb.2012.12.001. URL: http://dx.doi.org/10.1016/j.chb.2012.12.001.

Rosipal, Roman and Nicole Kr (2006). "Overview and Recent Advances in Partial Least Squares". In: *Subspace, Latent Structure and Feature Selection, Saunders, C., et al. (eds.) (Heidelberg: Springer-Verlag, 2006)*. ISSN: 03029743. DOI: 10.1007/11752790_2.

Russell, Stuart and Peter Norvig (2009). "Artificial Intelligence: A Modern Approach, 3rd edition". In: *Prentice Hall*, pp. 1–1132. ISSN: 0269-8889. DOI: 10.1017/S0269888900007724.

Ryan, R M and E L Deci (2000). "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being." In: *The American psychologist* 55.1, pp. 68–78. ISSN: 0003-066X. DOI: 10.1037/0003-066X.55.1.68.

Ryan, Richard M. et al. (Nov. 2006). "The Motivational Pull of Video Games: A Self-Determination Theory Approach". In: *Motivation and Emotion* 30.4, pp. 344–360. ISSN: 0146-7239. DOI: 10.1007/s11031-006-9051-8. URL: http://link.springer.com/10.1007/s11031-006-9051-8.

Sanders, Joshua I. et al. (2016). "Signatures of a Statistical Computation in the Human Sense of Confidence". In: *Neuron* 90.3, pp. 499–506. ISSN: 10974199. DOI: 10.1016/j.neuron.2016.03.025.

Schmidhuber, Jürgen (2010). "Formal theory of creativity, fun, and intrinsic motivation (1990–2010)". In: *IEEE Transactions on Autonomous Mental Development* 2.3, pp. 230–247.

Schmidt, Aaron M and Richard P DeShon (July 2007). "What to do? The effects of discrepancies, incentives, and time on dynamic goal prioritization." In: *The Journal of applied psychology* 92.4, pp. 928–41. ISSN: 0021-9010. DOI: 10.1037/0021-9010.92.4.928. URL: http://www.ncbi.nlm.nih.gov/pubmed/17638455.

Schmidt, Aaron M and Chad M Dolis (2009). "Something's got to give: The effects of dual-goal difficulty, goal progress, and expectancies on resource allocation." In: *Journal of Applied Psychology* 94.3, p. 678.

Schmidt, Kenneth A et al. (2010). "The ecology of information: an overview on the ecological significance of making informed decisions". In: *Oikos* 119.2, pp. 304–316.

Schüll, Natasha Dow and Ebook Library. (2012). *Addiction by design: machine gambling in Las Vegas*. xi, 442 p. ISBN: 978-0-691-16088-7. DOI: 10.2307/j.ctt12f4d0. URL: http://umichigan.eblib.com/patron/FullRecord.aspx?p=1023692%5Cnhttp://mirlyn.lib.umich.edu/Record/012927082%20CN%20-%20HV6721%20.L3%20S38%202012.

Seth, Anil K (Nov. 2013). *Interoceptive inference, emotion, and the embodied self*. Vol. 17. 11. Elsevier Ltd. 565–573. ISBN: 1879-307X (Electronic)$backslash$n1364-6613 (Linking). DOI: 10.1016/j.tics.2013.09.007. URL: http://www.ncbi.nlm.nih.gov/pubmed/24126130.

Sha, Lui et al. (2004). "Real Time Scheduling Theory: A Historical Perspective". In: *Real-Time Systems* 28.2, pp. 101–155. ISSN: 0922-6443. DOI: 10.1023/B:TIME.0000045315.61234.1e. URL: http://www.springerlink.com/openurl.asp?id=doi:10.1023/B:TIME.0000045315.61234.1e.

Shadmehr, Reza and Sandro Mussa-Ivaldi (2012). *Biological learning and control: how the brain builds representations, predicts events, and makes decisions*. MIT Press.

Shadmehr, Reza et al. (2016). "A Representation of Effort in Decision-Making and Motor Control". In: *Current Biology*. ISSN: 09609822. DOI: 10.1016/j.cub.2016.05.065.

Sheftel, Hila et al. (2013). "The geometry of the Pareto front in biological phenotype space". In: *Ecology and Evolution* 3.6, pp. 1471–1483. ISSN: 20457758. DOI: 10.1002/ece3.528.

Shenhav, Amitai et al. (2017). "Toward a rational and mechanistic account of mental effort". In: *Annual review of neuroscience* 40, pp. 99–124.

Shettleworth, S.J. (2010). *Cognition, evolution, and behavior*. Chicago: Oxford University Press.

Shiffrin, Richard M and Walter Schneider (1977). "Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory." In: *Psychological review* 84.2, p. 127.

Shin, Choonsung et al. (2012). "Understanding and Prediction of Mobile Application Usage for Smart Phones". In: *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. UbiComp '12. New York, NY, USA: ACM, pp. 173–182. ISBN: 978-1-4503-1224-0. DOI: 10.1145/2370216.2370243. URL: http://doi.acm.org/10.1145/2370216.2370243.

Silver, David et al. (2016). "Mastering the game of Go with deep neural networks and tree search". In: *Nature* 529.7587, pp. 484–489.

Simon, Herbert A (1967). "Motivational and emotional controls of cognition". In: *Psychological review* 74.1, pp. 29–39. ISSN: 0033-295X. DOI: http://dx.doi.org/10.1037/h0024127.

— (1994). "The bottleneck of attention: connecting thought with motivation." In: *In W. D. Spaulding (Ed.), Nebraska symposium on motivation, Vol. 41. Integrative views of motivation, cognition, and emotion*, pp. 1–21.

Simonov, Pavel V. (1984). *The need-informational theory of emotions*. Vol. 1. 3. 277–289. ISBN: 0167-8760 (Print)$backslash$r0167-8760 (Linking). DOI: 10.1016/0167-8760(84)90047-3.

Sinatra, Gale M et al. (2015). *The challenges of defining and measuring student engagement in science*. Taylor & Francis.

Sinha, Pawan et al. (2014). “Autism as a disorder of prediction”. In: *Proceedings of the National Academy of Sciences* 111.42, pp. 15220–15225.

Smith, EA et al. (1983). “Anthropological Applications of Optimal Foraging Theory: A Critical Review”. In: *Current ldots*. URL: http://www.jstor.org/stable/2743169.

Sperling, GDBA and BA Dosher (1986). “Handbook of perception and human performance”. In: *Handbook of perception and human performance. New York: Wiley* 1, pp. 1–65.

Srivastava, N and PR Schrater (2012). “Rational inference of relative preferences”. In: *Advances in Neural Information ldots*. URL: http://papers.nips.cc/paper/4670-rational-inference-of-relative-preferences.

Srivastava, Nisheeth and Paul Schrater (2015). “Learning What to Want: Context-Sensitive Preference Learning”. In: *PloS one* 10.10, e0141129.

Srivastava, Nisheeth et al. (2016). “Modeling sampling duration in decisions from experience”. In: *Proceedings of the 38th Annual Conference of the Cognitive Science Society*, pp. 2285–2290.

Stephens, D W and J R Krebs (1986). *Foraging Theory*. Vol. 121. 247 pp. ISBN: 0-691-08442-4. DOI: 10.2307/2409049. URL: http://books.google.com/books?hl=en&lr=&id=DVxvu-qDsaIC&pgis=1.

Stephens, David W (Dec. 2008). “Decision ecology: foraging and the ecology of animal decision making.” In: *Cognitive, affective & behavioral neuroscience* 8.4, pp. 475–84. ISSN: 1530-7026. DOI: 10.3758/CABN.8.4.475. URL: http://www.ncbi.nlm.nih.gov/pubmed/19033242.

Stephens, David W and Dack Anderson (2001). “The adaptive value of preference for immediacy : when shortsighted rules have farsighted consequences”. In: *Behavioral Ecology* 12, pp. 330–339. ISSN: 1045-2249, 1465-7279. DOI: 10.1093/beheco/12.3.330.

Stephens, David W and Eric L Charnov (1982). “Optimal foraging: some simple stochastic models”. In: *Behavioral Ecology and Sociobiology* 10.4, pp. 251–263.

Stephens, David W et al. (2007). *Foraging: Behavior and Ecology*. 608 pp. ISBN: 978-0-226-77264-6. URL: http://books.google.com/books?hl=en&lr=&id=5UvZT5BL6ZkC&oi=fnd&pg=PR5&dq=Foraging+behavior+and+ecology+(Stephens+DW+(Brown+JS+Ydenberg+RC))&ots=zEg1WGgAoM&sig=m3n9ombxBm1M-NYCAizmmmLFJ0I%255Cnpapers2://publication/uuid/74A7A495-BA10-4426-83C8-437F14B8E290%255.

Stoop, Paul P.M. and Vincent C.S. Wiers (1996). “The complexity of scheduling in practice”. In: *International Journal of Operations & Production Management* 16.10, pp. 37–53. ISSN: 0144-3577. DOI: 10.1108/01443579610130682.

Sukhbaatar, Sainbayar et al. (2018). “Learning Goal Embeddings via Self-Play for Hierarchical Reinforcement Learning”. In: *arXiv preprint arXiv:1811.09083*.

Sutton, Richard S and Andrew G Barto (1998). *Reinforcement Learning : An Introduction*. ISBN: 0-262-19398-1. DOI: 10.1016/j.brainres.2010.09.091.

Sutton, Richard S et al. (1999). "Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning". In: *Artificial intelligence* 112.1, pp. 181–211.

Tanaka, Saori C et al. (2004). "Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops". In: *Nature neuroscience* 7.8, pp. 887–893.

Terekhov, Daria et al. (July 2014). "Queueing-theoretic approaches for dynamic scheduling: A survey". In: *Surveys in Operations Research and Management Science* 19.2, pp. 105–129. ISSN: 18767354. DOI: 10.1016/j.sorms.2014.09.001. URL: http://linkinghub.elsevier.com/retrieve/pii/S1876735414000233.

Thompson, Edward P (1967). "Time, work-discipline, and industrial capitalism". In: *Past & Present* 38, pp. 56–97.

Thórisson, Kristinn R. et al. (2016). "Why artificial intelligence needs a task theory: And what it might look like". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. ISBN: 978-3-319-41648-9. DOI: 10.1007/978-3-319-41649-6_12.

Thura, D. et al. (2012). "Decision-making by urgency-gating: theory and experimental support". In: *Journal of Neurophysiology*, pp. 2912–2930. ISSN: 0022-3077. DOI: 10.1152/jn.01071.2011.

Tinbergen, N. (1963). *On aims and methods of Ethology*. Vol. 20. 410–433.

Tinbergen, Niko (1951). "The study of instinct." In:

Todorov, Emanuel and Michael I Jordan (2002). "Optimal feedback control as a theory of motor coordination". In: *Nature neuroscience* 5.11, p. 1226.

Todorov, Emanuel et al. (2005). "From task parameters to motor synergies: A hierarchical framework for approximately optimal control of redundant manipulators". In: *Journal of robotic systems* 22.11, pp. 691–710.

Tracy, J I et al. (2000). "The effect of autonomic arousal on attentional focus." In: *Neuroreport* 11.18, pp. 4037–4042. ISSN: 0959-4965. DOI: 10.1097/00001756-200012180-00027.

Trimmer, Pete C et al. (2013). "On the Evolution and Optimality of Mood States". In: *Behavioral Sciences* 3.3, pp. 501–521. ISSN: 2076-328X. DOI: 10.3390/bs3030501. URL: http://www.mdpi.com/2076-328X/3/3/501/htm.

Turner, Brandon M. et al. (2016). "Why more is better: Simultaneous modeling of EEG, fMRI, and behavioral data". In: *NeuroImage* 128, pp. 96–115. ISSN: 10959572. DOI: 10.1016/j.neuroimage.2015.12.030. URL: http://dx.doi.org/10.1016/j.neuroimage.2015.12.030.

Turner, Brandon M. et al. (2017). "Factor analysis linking functions for simultaneously modeling neural and behavioral data". In: *NeuroImage* 153 (July 2016), pp. 28–48. ISSN: 10959572. DOI: 10.1016/j.neuroimage.2017.03.044. URL: http://dx.doi.org/10.1016/j.neuroimage.2017.03.044.

Tversky, Amos and Daniel Kahneman (1973). "Availability: A heuristic for judging frequency and probability". In: *Cognitive psychology* 5.2, pp. 207–232.

Tversky, Amos and Daniel Kahneman (1986). "Rational Choice and the Framing of Decisions". In: *The Journal of Business* 59.4, pp. 251–278. ISSN: 0021-9398. DOI: 10.1086/296365.

United States. (2003). *American time use survey*. Washington, DC, US: U.S. Bureau of Labor Statistics. URL: http://purl.access.gpo.gov/GPO/LPS95587.

Vagni, Giacomo and Benjamin Cornwell (2018). "Patterns of everyday activities across social contexts". In: *Proceedings of the National Academy of Sciences* 115.24, pp. 6183–6188.

Van Der Schuur, Winneke A et al. (2015). "The consequences of media multitasking for youth: A review". In: *Computers in Human Behavior* 53, pp. 204–215.

Vancouver, Jeffrey B et al. (2010). "A formal, computational theory of multiple-goal pursuit: Integrating goal-choice and goal-striving processes." In: *Journal of Applied Psychology* 95.6, p. 985.

Vandekerckhove, Joachim (2014). "A cognitive latent variable model for the simultaneous analysis of behavioral and personality data". In: *Journal of Mathematical Psychology* 60, pp. 58–71. ISSN: 10960880. DOI: 10.1016/j.jmp.2014.06.004. URL: http://dx.doi.org/10.1016/j.jmp.2014.06.004.

Vezhnevets, Alexander Sasha et al. (2017). "Feudal networks for hierarchical reinforcement learning". In: *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. JMLR. org, pp. 3540–3549.

Vickers, Douglas et al. (1985). "Experimental paradigms emphasising state or process limitations: II effects on confidence". In: *Acta Psychologica*. ISSN: 00016918. DOI: 10.1016/0001-6918(85)90018-6.

Voss, Andreas et al. (2004). "Interpreting the parameters of the diffusion model: An empirical validation". In: *Memory and Cognition* 32.7, pp. 1206–1220. ISSN: 0090502X. DOI: 10.3758/BF03196893.

Vzquez, Alexei et al. (2006). "Modeling bursts and heavy tails in human dynamics". In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 73.3, pp. 1–19. ISSN: 15393755. DOI: 10.1103/PhysRevE.73.036127.

Wagenmakers, Eric Jan et al. (2007). "An EZ-diffusion model for response time and accuracy". In: *Psychonomic Bulletin and Review*. ISSN: 10699384. DOI: 10.3758/BF03194023.

Wald, A. (1945). "Sequential Tests of Statistical Hypotheses". In: *The Annals of Mathematical Statistics*. ISSN: 0003-4851. DOI: 10.1214/aoms/1177731118.

Walsh, Lynda (2013). *Scientists as prophets: A rhetorical genealogy*. Oxford University Press.

Wang, Zheng and John M. Tchernev (2012). "The "Myth" of Media Multitasking: Reciprocal Dynamics of Media Multitasking, Personal Needs, and Gratifications". In: *Journal of Communication* 62.3, pp. 493–513. ISSN: 00219916. DOI: 10.1111/j.1460-2466.2012.01641.x.

Wang, Zheng et al. (2015). “Multidimensions of media multitasking and adaptive media selection”. In: *Human Communication Research* 41.1, pp. 102–127.

Watts, Tyler W et al. (2018). “Revisiting the Marshmallow Test: A Conceptual Replication Investigating Links Between Early Delay of Gratification and Later Outcomes”. In: *Psychological science*, p. 0956797618761661.

Webb, Ian C. et al. (2009). “Bidirectional interactions between the circadian and reward systems: Is restricted food access a unique zeitgeber?” In: *European Journal of Neuroscience* 30.9, pp. 1739–1748. ISSN: 0953816X. DOI: 10.1111/j.1460-9568.2009.06966.x.

Weizenbaum examines computers and society (n.d.). URL: http://tech.mit.edu/V105/N16/weisen.16n.html.

White, Lawrence T et al. (2011). “What is the meaning of “on time”? The sociocultural nature of punctuality”. In: *Journal of Cross-Cultural Psychology* 42.3, pp. 482–493.

Whitton, Nicola and Alex Moseley (2014). “Deconstructing engagement: Rethinking involvement in learning”. In: *Simulation & Gaming* 45.4, pp. 433–449.

Wiecki, Thomas V et al. (2013). “HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python.” In: *Frontiers in neuroinformatics* 7, p. 14. ISSN: 1662-5196. DOI: 10.3389/fninf.2013.00014. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3731670&tool=pmcentrez&rendertype=abstract.

Wiener, Norbert (1961). *Cybernetics or Control and Communication in the Animal and the Machine*. Vol. 25. MIT press.

Winkielman, Piotr and Kent Berridge (2003). *Irrational Wanting and Subrational Liking: How Rudimentary Motivational and Affective Processes Shape Preferences and Choices*. Vol. 24. 4. 657–680. ISBN: 0162895X. DOI: 10.1046/j.1467-9221.2003.00346.x.

Wold, Svante et al. (2001). “PLS-regression: A basic tool of chemometrics”. In: *Chemometrics and Intelligent Laboratory Systems*. ISBN: 0169-7439. DOI: 10.1016/S0169-7439(01)00155-1.

Wrosch, Carsten et al. (2003a). “Adaptive self-regulation of unattainable goals: Goal disengagement, goal reengagement, and subjective well-being”. In: *Personality and Social Psychology Bulletin* 29.12, pp. 1494–1508.

Wrosch, Carsten et al. (2003b). “The importance of goal disengagement in adaptive self-regulation: When giving up is beneficial”. In: *Self and Identity* 2, pp. 1–20. URL: http://www.tandfonline.com/doi/abs/10.1080/15298860309021.

Xu, Zhiwei et al. (2007). “Effects of enriched environment on morphine-induced reward in mice”. In: *Experimental Neurology* 204, pp. 714–719. ISSN: 00144886. DOI: 10.1016/j.expneurol.2006.12.027.

Yeung, Nick and Christopher Summerfield (2012). “Metacognition in human decision-making: Confidence and error monitoring”. In: *Philosophical Transactions of the Royal*

Society B: Biological Sciences 367.1594, pp. 1310–1321. ISSN: 14712970. DOI: 10 . 1098/rstb.2011.0416.

Yu, Angela J. (2007). *Adaptive Behavior: Humans Act as Bayesian Learners*. ISBN: 0960-9822 (Print)$backslash$r0960-9822 (Linking). DOI: 10.1016/j.cub.2007.09.007.

Yu, Angela J. and Peter Dayan (2005). "Uncertainty, neuromodulation, and attention". In: *Neuron* 46.4, pp. 681–692. ISSN: 08966273. DOI: 10.1016/j.neuron.2005.04.026.