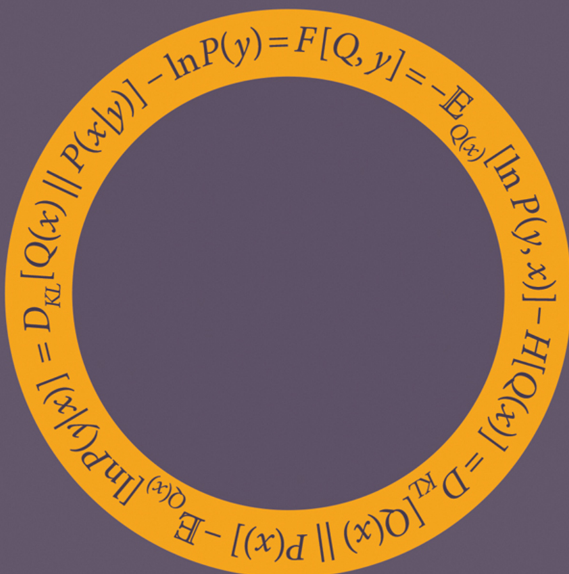
ACTIVE INFERENCE

The Free Energy Principle in Mind, Brain, and Behavior



THOMAS PARR

GIOVANNI PEZZULO

KARL J. FRISTON

"Probably the most lucid and comprehensive treatment of the concept of active inference to date."

—**Tomás Ryan**, Trinity College Dublin

Active Inference

Active Inference

The Free Energy Principle in Mind, Brain, and Behavior

Thomas Parr, Giovanni Pezzulo, and Karl J. Friston

The MIT Press
Cambridge, Massachusetts
London, England

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Parr, Thomas, 1993– author. | Pezzulo, Giovanni, author. | Friston, K. J. (Karl J.), author.
Title: Active inference : the free energy principle in mind, brain, and behavior / Thomas Parr, Giovanni Pezzulo, and Karl J. Friston.
Description: Cambridge, Massachusetts : The MIT Press, [2022] | Includes bibliographical references and index.
Identifiers: LCCN 2021023032 | ISBN 9780262045353 (hardcover)
Subjects: LCSH: Perception. | Inference. | Neurobiology. | Human behavior models. | Knowledge, Theory of. | Bayesian statistical decision theory.
Classification: LCC BF311 .P31366 2022 | DDC 153—dc23
LC record available at https://lccn.loc.gov/2021023032

Contents

Preface

Karl Friston

Active Inference is a way of understanding sentient behavior. The very fact that you are reading these lines means that you are engaging in Active Inference—namely, actively sampling the world—in a particular way—because you believe you will learn something. You are palpating this page with your eyes simply because this is the kind of action that will resolve uncertainty about what you will see next and—indeed—what these words convey. In short, Active Inference puts the action into perception, whereby perception is treated as perceptual inference or hypothesis testing. Active Inference goes even further and considers planning as inference—that is, inferring what you would do next to resolve uncertainty about your lived world.

To illustrate the simplicity of Active Inference—and what we are trying to explain—place your fingertips gently on your leg. Keep them there motionless for a second or two. Now, does your leg feel rough or smooth? If you had to move your fingers to evince a feeling of roughness or smoothness, you have discovered a fundament of Active Inference. To feel is to palpate. To see is to look. To hear is to listen. This palpation does not necessarily have to be overt—we can act covertly by directing our attention to this or that. In short, we are not simply trying to make sense of our sensations; we have to actively create our sensorium. In what follows, we will see why this has to be the case and why everything that we perceive, do, or plan is in the compass of one existential imperative—self-evidencing.

Active Inference is not just about reading or epistemic foraging. It is, on one view, something that all creatures and particles do, in virtue of their existence. This might sound like a strong claim; however, it speaks to the fact that Active Inference inherits from a free energy principle that equates existence with self-evidencing and self-evidencing with an enactive sort of inference. However, this book is not concerned with the physics of sentient systems. Its focus is on the implications of this physics for understanding how the brain works.

This understanding is not an easy business, as witnessed by millennia of natural philosophy and centuries of neuroscience. Although one can find the roots of Active Inference in first principle accounts of self-organized behavior (i.e., variational principles akin to Hamilton's principle of stationary action), first principles do not help very much when asking how a particular brain works and how it differs from another brain. For example, committing to the theory of evolution by natural selection does not help in the slightest when it comes to understanding why I have two eyes or speak French. This book is about using principles to scaffold key questions in neuroscience and artificial intelligence. To do this, we have to move beyond principles and get to grips with the mechanics to which the principles apply.

As such, Active Inference—and its accompanying Bayesian mechanics—is there to frame questions about how we perceive, plan, and act. Crucially, it does not aim to replace other frameworks, such as behavioral psychology, decision theory, and reinforcement learning. Rather, it hopes to embrace all those approaches that have proven so successful within a unified framework. In what follows, we will pay special attention to linking key constructs from psychology, cognitive neuroscience, enactivism, ethology, and so on to the calculus of belief updating in Active Inference—and its associated process theories.

By *process theories*, we refer to theories about how belief updating is realized by neuronal (and other biophysical) processes in the embodied brain and beyond. Work to date in Active Inference offers a fairly straightforward set of computational architectures and simulation tools to both model various aspects of a functioning brain and enable people to test hypotheses about different computational architectures. However, these tools only solve half the problem. At the heart of Active Inference lies a generative model—namely, a probabilistic representation of how unobservable causes in the world out there generate the observable consequences—our sensations. Getting the generative model right—as an apt explanation for the sentient behavior of any experimental subject or creature—is the big challenge.

This book tries to explain how to meet this challenge. The first part sets up the basic ideas and formalisms that are called on in the second part—to illustrate how they can be applied in practice. In short, this book is for people who want to use Active Inference to simulate and model sentient behavior, in the service of either scientific inquiry or, possibly, artificial intelligence. Thus it focuses on those ideas and procedures that are necessary to understand and implement an Active Inference scheme without getting distracted by the physics of sentient systems on the one hand or philosophy on the other.

A Note from Karl Friston

I have a confession to make. I did not write much of this book. Or, more precisely, I was not allowed to. This book's agenda calls for a crisp and clear writing style that is beyond me. Although I was allowed to slip in a few of my favorite words, what follows is a testament to Thomas and Giovanni, their deep understanding of the issues at hand, and, importantly, their theory of mind—in all senses.

Acknowledgments

We gratefully acknowledge invaluable input from our friends and colleagues—in particular, past and present members of the Theoretical Neurobiology group at the Wellcome Centre for Human Neuroimaging, University College London; the Cognition in Action (CONAN) Lab at the Institute of Cognitive Sciences and Technologies, National Research Council of Italy; and numerous international collaborators who have been integral to the development of the ideas presented in this book. This young but growing community has been more than generous in providing both intellectual support and motivation. Furthermore, we gratefully acknowledge Robert Prior and Anne-Marie Bono from MIT Press for kindly accompanying and advising us during the preparation of this book and Jakob Hohwy and other thoughtful reviewers for their guidance. Finally, we thank the funding agencies that provided financial support for our research: KJF was funded by a Wellcome Trust Principal Research Fellowship (Ref: 088130/Z/09/Z); GP was funded by the European Research Council under the Grant Agreement No. 820213 (ThinkAhead) and the European Union's Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement No. 945539 (Human Brain Project SGA3).

I

1 Overview

Chance favors the prepared mind.
—Louis Pasteur

1.1 Introduction

This chapter introduces the main question that Active Inference seeks to address: How do living organisms persist while engaging in adaptive exchanges with their environment? We discuss the motivation for addressing this question from a normative perspective, which starts from first principles and then unpacks their cognitive and biological implications. Furthermore, this chapter briefly introduces the structure of the book, including its subdivision into two parts: the first of which aims to help readers understand Active Inference, and the second of which aims to help them use it in their own research.

1.2 How Do Living Organisms Persist and Act Adaptively?

Living organisms constantly engage in reciprocal interactions with their environment (including other organisms). They emit *actions* that change the environment and receive *sensory observations* from it, as schematically illustrated in figure 1.1.

Living organisms can only maintain their bodily integrity by exerting adaptive *control* over the action-perception loop. This means acting to solicit sensory observations that either correspond to desired outcomes or goals (e.g., the sensations that accompany secure nutrients and shelter for simple

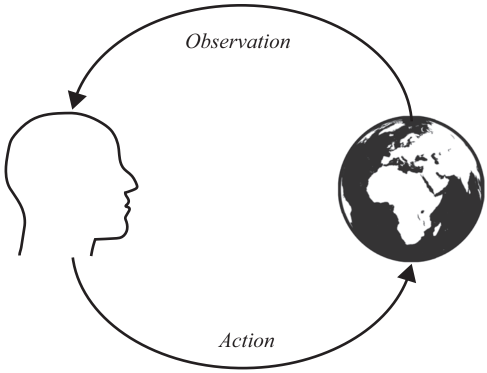


Figure 1.1
An action-perception cycle reciprocally connecting a creature and its environment. The term *environment* is intentionally generic. In the examples that we discuss, it can include the physical world, the body, the social environment, and so on.

organisms, or friends and jobs for more complex ones) or help in making sense of the world (e.g., informing the organism about its surroundings).

Engaging in adaptive action-perception loops with the environment poses formidable challenges to living organisms. This is largely due to the recursive nature of the cycle, where each observation, solicited by the previous action, changes how we decide on the next action, to solicit the next observation. The possibilities for control and adaptation are plentiful, but very few are useful. Yet during evolution, living organisms have managed to develop adaptive strategies to face the fundamental challenges of existence. These strategies vary in their level of cognitive sophistication, with simpler and more rigid solutions in simpler organisms (e.g., following nutrient gradients in bacteria) and more cognitively demanding and flexible solutions in more advanced organisms (e.g., planning to achieve distal goals in humans). These strategies also vary for the timescales at which they are selected and operate—ranging from simple responses to environmental threats or morphological adaptations that arise at an evolutionarily timescale, to behavioral patterns established during cultural or developmental learning, up to those requiring cognitive processes that operate at comparable timescales to action and perception (e.g., attention and memory).

1.3 Active Inference: Behavior from First Principles

This diversity is a blessing for biology but challenging for formal theories of brain and mind. Broadly, there are two perspectives we could take on this. One perspective is that different biological adaptations, neural processes (e.g., synaptic exchanges and brain networks), and cognitive mechanisms (e.g., perception, attention, social interaction) are highly idiosyncratic and require dedicated explanations. This would lead to proliferation of theories in fields like philosophy, psychology, neuroscience, ethology, biology, artificial intelligence, and robotics, with little hope for their unification. Another perspective is that, despite their diverse manifestations, the central aspects of behavior, cognition, and adaptation in living organisms are amenable to a coherent explanation from first principles.

These two possibilities map to two different research programs and, to some extent, different attitudes toward science: "neats" versus "scruffies" (terms due to Roger Shank). Neats always seek unification beyond the (apparent) heterogeneity of brain and mind phenomena. This usually corresponds to designing top-down, normative[1] models that start from first principles and try to derive as much as possible about brains and minds. Scruffies instead embrace the heterogeneity by focusing on details that demand dedicated explanations. This usually corresponds to designing bottom-up models that start from data and use whatever works to explain complex phenomena, including different explanations for different phenomena.

Is it possible to explain heterogenous biological and cognitive phenomena from first principles, as the neats assume? Is a unified framework to understand brain and mind possible?

This book answers these questions affirmatively and advances Active Inference as a normative approach to understand brain and mind. Our treatment of Active Inference starts from first principles and unpacks their cognitive and biological implications.

1.4 Structure of the Book

The book comprises two parts. These are aimed at readers who want to understand Active Inference (first part) and those who seek to use it for their own research (second part). The first part of the book introduces Active Inference both conceptually and formally, contextualizing it within current theories of cognition. The goal of this first part is to provide a comprehensive, formal, and self-contained introduction to Active Inference: its main constructs and implications for the study of brain and cognition.

The second part of the book illustrates specific examples of computational models that use Active Inference to explain cognitive phenomena, such as perception, attention, memory, and planning. The goal of this second part is to help readers both understand existing computational models using Active Inference and design novel ones. In short, this book divides into theory (part 1) and practice (part 2).

1.4.1 Part 1: Active Inference in Theory

Active Inference is a normative framework to characterize Bayes-optimal[2] behavior and cognition in living organisms. Its normative character is evinced in the idea that all facets of behavior and cognition in living organisms follow a unique imperative: *minimizing the surprise of their sensory observations*. *Surprise* has to be interpreted in a technical sense: it measures how much an agent's current sensory observations differ from its preferred sensory observations—that is, those that preserve its integrity (e.g., for a fish, being in the water). Importantly, minimizing surprise is not something that can be done by passively observing the environment: rather, agents must adaptively *control* their action-perception loops to solicit desired sensory observations. This is the active bit of Active Inference.

Minimizing surprise turns out to be a challenging problem for technical reasons that will become apparent later. Active Inference offers a solution to this problem. It assumes that even if living organisms cannot directly minimize their surprise, they can minimize a proxy—called *(variational) free energy*. This quantity can be minimized through neural computation in response to (and in anticipation of) sensory observations. This emphasis on free energy minimization discloses the relation between Active Inference and the (first) principle that motivates it: the *free energy principle* (Friston 2009).

Free energy minimization seems a very abstract starting point to explain biological phenomena. However, it is possible to derive a number of formal and empirical implications from it and to address a number of central questions in cognitive and neural theory. These include how the variables involved in free energy minimization may be encoded in neuronal populations; how the computations of minimized free energy map to specific cognitive processes, such as perception, action selection, and learning; and what kind of behaviors emerge when an Active Inference agent minimizes its free energy.

As the above list of topics exemplifies, in this book we are mainly concerned with Active Inference and free energy minimization at the level of living organisms—simpler (e.g., bacterial) or more complex (e.g., human)—and their behavioral, cognitive, social, and neural processes. This clarification is necessary to contextualize our treatment of Active Inference within the more general free energy principle (FEP), which discusses free energy minimization across a much wider range of biological phenomena and timescales beyond neural information processing—ranging from evolutionary to cellular and cultural (Friston, Levin et al. 2015; Isomura and Friston 2018; Palacios, Razi et al. 2020; Veissière et al. 2020)—which are beyond the scope of this book.

It is possible to motivate Active Inference by taking one of two roads: a high road and a low road; see figure 1.2. These two roads provide two distinct but highly complementary perspectives on Active Inference:

- The *high road* to Active Inference starts from the question of how living organisms persist and act adaptively in the world and motivates

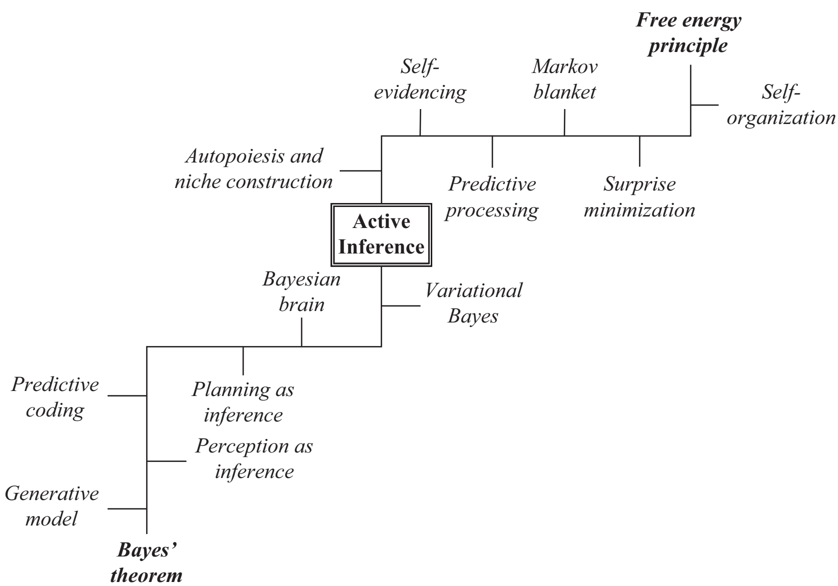


Figure 1.2
Two roads to Active Inference: the high road (starting from top-right) and the low road (starting from bottom-left).

Active Inference as a normative solution to these problems. This high road perspective is useful to understand the normative nature of Active Inference: *what* living organisms must do to face their fundamental existential challenges (minimize their free energy) and *why* (to vicariously minimize the surprise of their sensory observations).

- The *low road* to Active Inference starts from the notion of the Bayesian brain, which casts the brain as an inference engine trying to optimize probabilistic representations of the causes of its sensory input. It then motivates Active Inference as a specific, variational approximation to the (otherwise intractable) inferential problem, which has a degree of biological plausibility. This low road perspective is useful to illustrate *how* Active Inference agents minimize their free energy—therefore illustrating Active Inference not just as a principle but also as a mechanistic explanation (aka *process theory*) of cognitive functions and their neuronal underpinnings.

In chapter 2, we set out the low road perspective on Active Inference. We start from foundational theories that cast perception as a problem of statistical (Bayesian) inference (Helmholtz 1866) and their modern incarnation in the Bayesian brain hypothesis (Doya 2007). We will see that to perform such (perceptual) inference, living organisms must be equipped with—or embody—a *probabilistic generative model* of how their sensory observations are generated, which encodes beliefs (probability distributions) about both observable variables (sensory observations) and nonobservable (hidden) variables. We will extend this inferential view beyond perception to cover problems of action selection, planning, and learning.

In chapter 3, we will illustrate the complementary high road perspective on Active Inference. This chapter introduces the FEP and the imperative for biological organisms to minimize surprise. Further to this, it unpacks how this principle encompasses the dynamics of self-organization and the preservation of a statistical boundary or *Markov blanket* that maintains separation from the environment. This is vital in maintaining the integrity of biological creatures, and it is central to their autopoiesis.

In chapter 4, we will unpack Active Inference more formally. This chapter takes its cue from the discussion of the Bayesian brain in chapter 2 and sets out the mathematical relationship between the self-evidencing dynamics of chapter 3 and variational inference. In addition, this chapter sets out two sorts of generative model used to formulate Active Inference problems. These include the partially observed Markov decision processes used for decision-making and planning and the continuous time dynamical models that interface with sensory receptors and muscles. Finally, we see how free energy minimization for each of these models manifests as dynamic belief updating.

In chapter 5, we will move from formal treatments to biological implications of Active Inference. By starting from the premise that "everything that changes in the brain must minimize free energy" (Friston 2009), we will discuss how the specific quantities involved in the free energy minimization (e.g., prediction, prediction error, and precision signals) manifest in neuronal dynamics. This aids in mapping the abstract computational principles of Active Inference to specific neural computations that can be executed by physiological substrates. This is important in forming hypotheses under this framework and ensures that these are answerable to measured data. In other words, chapter 5 sets out the process theory associated with Active Inference.

Throughout the first part of the book, we will discuss several characteristic aspects of Active Inference. These highlight the ways in which it is different from alternative frameworks that seek to explain biological regulation and cognition—some of which we preview here.

- Under Active Inference, *perception* and *action* are two complementary ways to fulfill the same imperative: minimization of free energy. Perception minimizes free energy (and surprise) by (Bayesian) belief updating or *changing your mind*, thus making your beliefs compatible with sensory observations. Instead, action minimizes free energy (and surprise) by *changing the world* to make it more compatible with your beliefs and goals. This unification of cognitive functions marks a fundamental difference between Active Inference and other approaches that treat action and perception in isolation from one another. Learning is yet another way to minimize free energy. However, it is not fundamentally different from perception; it simply operates at a slower timescale. The complementarity between perception and action will be unpacked in chapter 2.
- In addition to driving action selection in the present to change currently available sensory data, the Active Inference framework accommodates planning—or the selection of the optimal course of action (or policy) in the future. Optimality here is measured in relation to an *expected free*

energy and is distinct from the notion of *variational free energy* considered above in the context of action and perception. Indeed, while computing variational free energy depends on present and past observations, computing expected free energy also requires predicted future observations (hence the term *expected*). Interestingly, the expected free energy of a policy comprises two parts. The first quantifies the extent to which the policy is expected to resolve uncertainty (exploration) and the second how consistent the predicted outcomes are with an agent's goals (exploitation). In contrast with other frameworks, policy selection in Active Inference automatically balances exploration and exploitation. The relations between variational and expected free energy will be unpacked in chapter 2.

- Under Active Inference, all cognitive operations are conceptualized as inference over *generative models*—in keeping with the idea that the brain performs probabilistic computations—aka *the Bayesian brain hypothesis*. Yet, the appeal to a specific approximate form of Bayesian inference—that is, a variational scheme that is motivated by first principles—adds specificity to the process theory. Furthermore, Active Inference extends the inferential approach to domains of cognition that are rarely considered and adds some specificity to the kind of models and inferential processes that may be implemented by biological brains. Under some assumptions, the dynamics that emerge from generative models used in Active Inference closely correspond to widespread models in computational neuroscience, such as predictive coding (Rao and Ballard 1999) and the Helmholtz machine (Dayan et al. 1995). The specifics of the variational scheme will be unpacked in chapter 4.
- Under Active Inference, both perception and learning are *active* processes, for two reasons. First, the brain is essentially a *predictive machine*, which constantly predicts incoming stimuli rather than passively waiting for them. This is important as perceptual and learning processes are always contextualized by prior predictions (e.g., expected and unexpected stimuli affect perception and learning in different ways). Second, creatures engaging in Active Inference actively seek out *salient* sensory observations that resolve their uncertainty (e.g., by orienting their sensors or selecting learning episodes that are informative). The active character of perception and learning stands in contrast with most current theories that treat them as largely passive processes; this will be unpacked in chapter 2.

- Action is quintessentially goal directed and purposive. It starts from a desired outcome or goal (analogous to the concept of a set-point in cybernetics), which is encoded as a prior prediction. Planning proceeds by inferring an action sequence that fulfills this prediction (or equivalently, reduces any prediction error between prior prediction and the current state). The goal-directed character of action in Active Inference is in keeping with early cybernetic formulations but is distinct from most current theories that explain behavior in terms of stimulus-response mappings or state-action policies. Stimulus-response or habitual behavior then becomes a special case of a broader family of policies in Active Inference. The goal-directed nature of Active Inference will be unpacked in chapters 2 and 3.
- Various constructs of Active Inference have plausible biological analogues in the brain. This implies that—once one has defined a specific generative model for a problem at hand—one can move from Active Inference as a normative theory to Active Inference as a process theory, which makes specific empirical predictions. For example, perceptual inference and learning correspond to changing synaptic activity and changing synaptic efficacy, respectively. Precision of predictions (in predictive coding) corresponds to the synaptic gain of prediction error units. Precision of policies corresponds to dopaminergic activity. Some of the biological consequences of Active Inference will be unpacked in chapter 5.

1.4.2 Part 2: Active Inference in Practice

While the first part of the book provides readers with the conceptual and formal tools to understand Active Inference, the second part focuses on practical issues. Specifically, we hope to provide readers with the tools to understand existing Active Inference models of cognitive functions (and dysfunctions) and to design novel ones. To this aim, we discuss specific examples of models using Active Inference. Importantly, models of Active Inference can vary along different dimensions (e.g., with discrete or continuous time formulations, flat or hierarchical inference). The second part is structured as follows:

In chapter 6, we introduce a recipe to build Active Inference models. The recipe covers the essential steps to design an effective model, which include the identification of the system of interest, the most appropriate form of the generative model (e.g., to characterize discrete- or continuous-time phenomena), and the specific variables to be included in the model. This chapter therefore offers an introduction to the design principles that underwrite the models discussed in the following chapters.

In chapter 7, we discuss Active Inference models that address problems formulated in discrete time; for example, as hidden Markov models (HMMs) or partially observable Markov decision processes (POMDPs). Our examples include a model of perceptual processing and a model of discrete foraging choices—that is, whether to turn left or right at a decision point to secure a reward. We also introduce topics such as information seeking, learning, and novelty seeking, which can be treated in terms of discrete-time Active Inference.

In chapter 8, we discuss Active Inference models that address problems formulated in continuous time, using stochastic differential equations. These include models of perception (like predictive coding), movement control, and sequential dynamics. Interestingly, it is in the continuous-time formulation that some of the most distinctive predictions of Active Inference appear, such as the idea that movement generation stems from the fulfillment of predictions and that attentional phenomena can be understood in terms of precision control. We also introduce hybrid models of Active Inference that include both discrete- and continuous-time variables. These permit simultaneous assessment of the choice among discrete options (e.g., targets for saccades) and the continuous movements resulting from the choice (e.g., oculomotor movements).

In chapter 9, we illustrate how to use Active Inference models to analyze data from behavioral experiments. We discuss the specific steps that are necessary for model-based data analysis, from the collection of data to the formulation of a model and its inversion to support the analysis of data from single participants or at the group level.

In chapter 10, we discuss the relations between Active Inference and other theories in psychology, neuroscience, AI, and philosophy. We also highlight the most important aspects of Active Inference that distinguish it from the other theories.

In the appendixes, we briefly discuss the mathematical background required to understand the most technical parts of the book, including the notions of Taylor series approximation, variational Laplace, variational calculus, and more. For reference we also present in a concise form the most important equations used in Active Inference.

In sum, the second part of the book illustrates a broad variety of models of biological and cognitive phenomena that can be constructed using Active Inference and a methodology to design novel ones. Apart from the interest of the specific models, we hope that our treatment clarifies the value of using a unified, normative framework to address biological and cognitive phenomena from a coherent perspective. In the end, this is the real appeal of normative frameworks: to provide a unified perspective and a guiding principle to reconcile apparently disconnected phenomena—in this case, phenomena like perception, decision-making, attention, learning, and movement control, each having its separate chapter in any psychology or neuroscience manual.

The models highlighted in the second part have been selected to illustrate specific points as simply as possible. While we cover several models and domains, from discrete-time decisions to continuous-time perception and movement control, we are clearly disregarding many others that are equally interesting. Many other Active Inference models exist in the literature that cover domains as diverse as biological self-organization and the origins of life (Friston 2013), morphogenesis (Friston, Levin et al. 2015), cognitive robotics (Pio-Lopez et al. 2016, Sancaktar et al. 2020), social dynamics and niche construction (Bruineberg, Rietveld et al. 2018), the dynamics of synaptic networks (Palacios, Isomura et al. 2019), learning in biological networks (Friston and Herreros 2016), and psychopathological conditions, such as post-traumatic stress disorder (Linson et al. 2020) and panic disorder (Maisto, Barca et al. 2021). These models vary along many dimensions: some are more directly related to biology whereas others are less so; some are single-agent models whereas others are multi-agent models; some target adaptive inference whereas other target maladaptive inference (e.g., in patient groups), and so on.

This growing literature exemplifies the increasing popularity of Active Inference and the possibility of using it in a very large variety of domains. The aim of this book is to provide our readers with the ability to understand and use Active inference in their own research—possibly, to explore its unforeseen potentialities.

1.5 Summary

This chapter briefly introduces the Active Inference approach to explain biological problems from a normative perspective—and previews some implications of this perspective that will be unpacked in later chapters. Furthermore, this chapter highlights the division of the book into two parts, which aim to help readers understand Active Inference and use it in their own research, respectively. Over the next few chapters, we will develop the low road and high road perspectives outlined herein, before delving into the structure of generative models and the resulting message passing. Together these comprise Active Inference in principle and provide the preliminaries for Active Inference in practice. We hope that these chapters will persuade readers that Active Inference offers not only a unifying principle under which to understand behavior but also a tractable approach to studying action and perception in autonomous systems.

References

Ackley, D. H., G. E. Hinton, and T. J. Sejnowski (1985). "A learning algorithm for Boltzmann machines." *Cognitive Science* 9(1): 147–169.

Adams, R. A., E. Aponte, L. Marshall, and K. J. Friston (2015). "Active inference and oculomotor pursuit: the dynamic causal modelling of eye movements." *Journal of Neuroscience Methods* 242: 1–14.

Adams, R. A., M. Bauer, D. Pinotsis, and K. J. Friston (2016). "Dynamic causal modelling of eye movements during pursuit: confirming precision-encoding in V1 using MEG." *NeuroImage* 132: 175–189.

Adams, R. A., L. U. Perrinet, and K. Friston (2012). "Smooth pursuit and visual occlusion: Active Inference and oculomotor control in schizophrenia." *PLOS ONE* 7(10): e47502.

Adams, R. A., S. Shipp, and K. J. Friston (2013). "Predictions not commands: Active Inference in the motor system." *Brain Structure and Function* 218(3): 611–643.

Adams, R. A., K. E. Stephan, H. R. Brown, C. D. Frith, and K. J. Friston (2013). "The computational anatomy of psychosis." *Frontiers in Psychiatry* 4: 47.

Aghajanian, G. K., and G. J. Marek (1999). "Serotonin, via 5-HT2A receptors, increases EPSCs in layer V pyramidal cells of prefrontal cortex by an asynchronous mode of glutamate release." *Brain Research* 825(1): 161–171.

Ahmadi, A., and J. Tani (2019). "A novel predictive-coding-inspired variational RNN model for online prediction and recognition." *Neural Computation* 31(11): 2025–2074.

Aitchison, L., and M. Lengyel (2017). "With or without you: predictive coding and Bayesian inference in the brain." *Current Opinion in Neurobiology* 46: 219–227.

Allen, M., A. Levy, T. Parr, and K. J. Friston (2019). "In the body's eye: the computational anatomy of interoceptive inference." *bioRxiv* 603928.

Anderson, B. A., P. A. Laurent, and S. Yantis (2011). "Value-driven attentional capture." *Proceedings of the National Academy of Sciences* 108(25): 10367.

Arnal, L. H., and A.-L. Giraud (2012). "Cortical oscillations and sensory predictions." *Trends in Cognitive Sciences* 16(7): 390–398.

Arnsten, A. F. T., and B.-M. Li (2005). "Neurobiology of executive functions: catecholamine influences on prefrontal cortical functions." *Biological Psychiatry* 57(11): 1377–1384.

Ashby, W. R. (1952). *Design for a Brain*. Oxford: Wiley.

Attias, H. (2003). "Planning by probabilistic inference." *Proceedings of the 9th International Workshop on Artificial Intelligence and Statistics, Key West, Florida, USA*.

Baldassarre, G., and M. Mirolli (2013). *Intrinsically Motivated Learning in Natural and Artificial Systems*. New York: Springer.

Barca, L., and G. Pezzulo (2020). "Keep your interoceptive streams under control: an Active Inference perspective on anorexia nervosa." *Cognitive, Affective and Behavioral Neuroscience* 20(2): 427–440.

Barrett, L. F. (2017). *How Emotions Are Made: The Secret Life of the Brain*. Boston, MA: Houghton Mifflin Harcourt.

Barrett, L. F., K. S. Quigley, and P. Hamilton (2016). "An Active Inference theory of allostasis and interoception in depression." *Philosophical Transactions of the Royal Society B* 371(1708): 20160011.

Barrett, L. F., and W. K. Simmons (2015). "Interoceptive predictions in the brain." *Nature Reviews Neuroscience* 16(7): 419–429.

Barsalou, L. W. (2008). "Grounded cognition." *Annual Review of Psychology* 59: 617–645.

Bastos, A. M., V. Litvak, R. Moran, C. A. Bosman, P. Fries, and K. J. Friston (2015). "A DCM study of spectral asymmetries in feedforward and feedback connections between visual areas V1 and V4 in the monkey." *Neuroimage* 108: 460–475.

Bastos, A. M., W. M. Usrey, R. A. Adams, G. R. Mangun, P. Fries, and K. J. Friston (2012). "Canonical microcircuits for predictive coding." *Neuron* 76(4): 695–711.

Beal, M. J. (2003). "Variational algorithms for approximate Bayesian inference." PhD diss., University of London.

Bellman, R. (1954). "The theory of dynamic programming." *Bulletin of the American Mathematical Society* 60(6): 503–515.

Benrimoh, D., T. Parr, P. Vincent, R. A. Adams, and K. Friston (2018). "Active Inference and auditory hallucinations." *Computational Psychiatry* 2: 183–204.

Berridge, K. C. (2007). "The debate over dopamine's role in reward: the case for incentive salience." *Psychopharmacology* 191(3): 391–431.

Berridge, K. C., and M. L. Kringelbach (2011). "Building a neuroscience of pleasure and well-being." *Psychology of Well-Being* 1(1): 1–3.

Botvinick, M., and M. Toussaint (2012). "Planning as inference." *Trends in Cognitive Sciences* 16(10): 485–488.

Botvinick, M. M. (2008). "Hierarchical models of behavior and prefrontal function." *Trends in Cognitive Sciences* 12(5): 201–208.

Brown, H., R. A. Adams, I. Parees, M. Edwards, and K. Friston (2013). "Active Inference, sensory attenuation and illusions." *Cognitive Processing* 14(4): 411–427.

Brown, H., and K. Friston (2012). "Free-energy and illusions: the cornsweet effect." *Frontiers in Psychology* 3(43).

Brown, L. D. (1981). "A complete class theorem for statistical problems with finite-sample spaces." *Annals of Statistics* 9(6): 1289–1300.

Bruineberg, J., J. Kiverstein, and E. Rietveld (2016). "The anticipating brain is not a scientist: the free-energy principle from an ecological-enactive perspective." *Synthese* 195: 2417–2444.

Bruineberg, J., E. Rietveld, T. Parr, L. van Maanen, and K. J. Friston (2018). "Free-energy minimization in joint agent-environment systems: a niche construction perspective." *Journal of Theoretical Biology* 455: 161–178.

Buzsaki, G. (2019). *The Brain from Inside Out*. New York: Oxford University Press.

Callaway, E. M., and A. K. Wiser (2009). "Contributions of individual layer 2–5 spiny neurons to local circuits in macaque primary visual cortex." *Visual Neuroscience* 13(5): 907–922.

Cannon, W. B. (1929). "Organization for physiological homeostasis." *Physiological Reviews* 9(3): 399–431.

Ciria, A., G. Schillaci, G. Pezzulo, V. V. Hafner, and B. Lara (2021). "Predictive processing in cognitive robotics: a review." *arXiv preprint arXiv:2101.06611*.

Cisek, P. (2019). "Resynthesizing behavior through phylogenetic refinement." *Attention, Perception, and Psychophysics* 81(7): 2265–2287.

Clark, A. (2013). "Whatever next? Predictive brains, situated agents, and the future of cognitive science." *Behavioral and Brain Sciences* 36(03): 181–204.

Clark, A. (2015). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. New York: Oxford University Press.

Clark, A., and D. J. Chalmers (1998). "The extended mind." *Analysis* 58: 10–23.

Clark, J. E., S. Watson, and K. J. Friston (2018). "What is mood? A computational perspective." *Psychological Medicine* 48(14): 2277–2284.

Collins, S. H., M. Wisse, and A. Ruina (2016). "A three-dimensional passive-dynamic walking robot with two legs and knees." *International Journal of Robotics Research* 20(7): 607–615.

Conant, R. C., and W. R. Ashby (1970). "Every good regulator of a system must be a model of that system." *International Journal of Systems Science* 1(2): 89–97.

Corcoran, A. W., G. Pezzulo, and J. Hohwy (2020). "From allostatic agents to counterfactual cognisers: Active Inference, biological regulation, and the origins of cognition." *Biology and Philosophy* 35(3): 32.

Corlett, P. R., G. Horga, P. C. Fletcher, B. Alderson-Day, K. Schmack, and A. R. Powers III (2019). "Hallucinations and strong priors." *Trends in Cognitive Sciences* 23(2): 114–127.

Cox, D. R., and H. D. Miller (1965). "The theory of stochastic processes." London: Chapman and Hall/CRC.

Craik, K. (1943). *The Nature of Explanation*. Cambridge: Cambridge University Press.

Cullen, M., B. Davey, K. J. Friston, and R. J. Moran (2018). "Active inference in OpenAI gym: a paradigm for computational investigations into psychiatric illness." *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 3(9): 809–818.

Da Costa, L., T. Parr, N. Sajid, S. Veselic, V. Neacsu, and K. Friston (2020). "Active Inference on discrete state-spaces: a synthesis." *arXiv preprint arXiv:2001.07203*.

Daunizeau, J., H. E. M. den Ouden, M. Pessiglione, S. J. Kiebel, K. E. Stephan, and K. J. Friston (2010). "Observing the observer (I): meta-Bayesian models of learning and decision-making." *PLOS ONE* 5(12): e15554.

Dauwels, J. (2007). "On variational message passing on factor graphs." *2007 IEEE International Symposium on Information Theory*, 2546–2550.

Daw, N. D., Y. Niv, and P. Dayan (2005). "Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control." *Nature Neuroscience* 8(12): 1704–1711.

Dayan, P., G. E. Hinton, R. M. Neal, and R. S. Zemel (1995). "The Helmholtz machine." *Neural Computation* 7: 889–904.

Demirdjian, D., L. Taycher, G. Shakhnarovich, K. Grauman, and T. Darrell (2005). "Avoiding the 'streetlight effect': tracking by exploring likelihood modes." In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, 357–364.

Dennett, D. C. (1978). "Why not the whole iguana?" *Behavioral and Brian Sciences* 1: 103–104.

Dickinson, A., and B. Balleine (1990). "Motivational control of instrumental performance following a shift from thirst to hunger." *Quarterly Journal of Experimental Psychology* 42(4): 413–431.

Disney, A. A., C. Aoki, and M. J. Hawken (2007). "Gain modulation by nicotine in Macaque V1." *Neuron* 56(4): 701–713.

Donnarumma, F., M. Costantini, E. Ambrosini, K. Friston, and G. Pezzulo (2017). "Action perception as hypothesis testing." *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior* 89: 45–60.

Doya, K. (2007). *Bayesian Brain: Probabilistic Approaches to Neural Coding*. Cambridge, MA: MIT Press.

Elliott, M. C., P. M. Tanaka, R. W. Schwark, and R. Andrade (2018). "Serotonin differentially regulates L5 pyramidal cell classes of the medial prefrontal cortex in rats and mice." *eNeuro* 5(1): eneuro.0305–0317.2018.

Feldman, A. G. (2009). "New insights into action-perception coupling." *Experimental Brain Research* 194(1): 39–58.

Feldman, A. G., and M. F. Levin (2009). "The equilibrium-point hypothesis—past, present and future." In *Progress in Motor Control: A Multidisciplinary Perspective*, edited by D. Sternad, 699–726. Boston, MA: Springer US.

Feldman, H., and K. Friston (2010). "Attention, uncertainty, and free-energy." *Frontiers in Human Neuroscience* 4(215).

Felleman, D. J., and D. C. Van Essen (1991). "Distributed hierarchical processing in the primate cerebral cortex." *Cerebral Cortex* 1(1): 1–47.

Fiser, J., P. Berkes, G. Orbán, and M. Lengyel (2010). "Statistically optimal perception and learning: from behavior to neural representations." *Trends in Cognitive Sciences* 14(3): 119–130.

FitzGerald, T. H. B., R. J. Dolan, and K. Friston (2015). "Dopamine, reward learning, and active inference." *Frontiers in Computational Neuroscience* 9: 1–16.

FitzGerald, T. H. B., P. Schwartenbeck, M. Moutoussis, R. J. Dolan, and K. Friston (2015). "Active inference, evidence accumulation, and the urn task." *Neural Computation* 27(2): 306–328.

Foster, D. (2019). *Generative Deep Learning: Teaching Machines to Paint, Write, Compose, and Play*. Boston: O'Reilly Media.

Fountas, Z., N. Sajid, P. A. M. Mediano, and K. Friston (2020). "Deep active inference agents using Monte-Carlo methods." *arXiv:2006.04176 [cs, q-bio, stat].*

Fradkin, I., R. A. Adams, T. Parr, J. P. Roiser, and J. D. Huppert (2020). "Searching for an anchor in an unpredictable world: a computational model of obsessive compulsive disorder." *Psychological Review* 127(5): 672–699.

Frank, M. J. (2005). "Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism." *Journal of Cognitive Neuroscience* 17(1): 51–72.

Freeze, B. S., A. V. Kravitz, N. Hammack, J. D. Berke, and A. C. Kreitzer (2013). "Control of basal ganglia output by direct and indirect pathway projection neurons." *Journal of Neuroscience* 33(47): 18531–18539.

Freund, T. F., J. F. Powell, and A. D. Smith (1984). "Tyrosine hydroxylase-immunoreactive boutons in synaptic contact with identified striatonigral neurons, with particular reference to dendritic spines." *Neuroscience* 13(4): 1189–1215.

Friston, K. (2005). "A theory of cortical responses." *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 360(1456): 815–836.

Friston, K. (2008). "Hierarchical models in the brain." *PLOS Computational Biology* 4(11): e1000211.

Friston, K. (2009). "The free-energy principle: a rough guide to the brain?" *Trends in Cognitive Sciences* 13(7): 293–301.

Friston, K. (2011). "What is optimal about motor control?" *Neuron* 72(3): 488–498.

Friston, K. (2013). "Life as we know it." *Journal of the Royal Society Interface* 10(86): 20130475.

Friston, K. (2017). "Precision psychiatry." *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 2(8): 640–643.

Friston, K. (2019a). "A free energy principle for a particular physics." *arXiv preprint arXiv:1906.10184*.

Friston, K. (2019b). "Waves of prediction." *PLOS Biology* 17(10): e3000426.

Friston, K., R. Adams, L. Perrinet, and M. Breakspear (2012). "Perceptions as hypotheses: saccades as experiments." *Frontiers in Psychology* 3(151).

Friston, K., and G. Buzsaki (2016). "The functional anatomy of time: what and when in the brain." *Trends in Cognitive Sciences* 20(7): 500–511.

Friston, K., L. Da Costa, D. Hafner, C. Hesp, and T. Parr (2020). "Sophisticated inference." *arXiv preprint arXiv:2006.04120*.

Friston, K., J. Daunizeau, and S. J. Kiebel (2009). "Reinforcement learning or Active Inference?" *PLOS ONE* 4(7): e6421.

Friston, K., J. Daunizeau, J. Kilner, and S. J. Kiebel (2010). "Action and behavior: a free-energy formulation." *Biological Cybernetics* 102(3): 227–260.

Friston, K., T. FitzGerald, F. Rigoli, P. Schwartenbeck, J. O'Doherty, and G. Pezzulo (2016). "Active Inference and learning." *Neuroscience and Biobehavioral Reviews* 68: 862–879.

Friston, K., T. FitzGerald, F. Rigoli, P. Schwartenbeck, and G. Pezzulo (2017). "Active Inference: a process theory." *Neural Computation* 29(1): 1–49.

Friston, K., and C. D. Frith (2015a). "Active inference, communication and hermeneutics( )." *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior* 68: 129–143.

Friston, K., and C. Frith (2015b). "A duet for one." *Consciousness and Cognition* 36: 390–405.

Friston, K., and I. Herreros (2016). "Active Inference and learning in the cerebellum." *Neural Computation* 28(9): 1812–1839.

Friston, K., and S. Kiebel (2009). "Predictive coding under the free-energy principle." *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1521): 1211.

Friston, K., M. Levin, B. Sengupta, and G. Pezzulo (2015). "Knowing one's place: a free-energy approach to pattern regulation." *Journal of the Royal Society Interface* 12(105): 20141383.

Friston, K., M. Lin, C. D. Frith, G. Pezzulo, J. A. Hobson, and S. Ondobaka (2017). "Active Inference, curiosity and insight." *Neural Computation* 29(10): 2633–2683.

Friston, K., V. Litvak, A. Oswal, A. Razi, K. E. Stephan, B. C. M. van Wijk, G. Ziegler, and P. Zeidman (2016). "Bayesian model reduction and empirical Bayes for group (DCM) studies." *NeuroImage* 128(Supplement C): 413–431.

Friston, K., J. Mattout, and J. Kilner (2011). "Action understanding and active inference." *Biological Cybernetics* 104(1): 137–160.

Friston, K., J. Mattout, N. Trujillo-Barreto, J. Ashburner, and W. Penny (2007). "Variational free energy and the Laplace approximation." *NeuroImage* 34(1): 220–234.

Friston, K., T. Parr, and B. de Vries (2017). "The graphical brain: belief propagation and Active Inference." *Network Neuroscience* 1(4): 381–414.

Friston, K., T. Parr, Y. Yufik, N. Sajid, C. J. Price, and E. Holmes (2020). "Generative models, linguistic communication and active inference." *Neuroscience and Biobehavioral Reviews* 118: 42–64.

Friston, K., T. Parr, and P. Zeidman (2018). "Bayesian model reduction." *arXiv preprint arXiv:1805.07092*.

Friston, K., F. Rigoli, D. Ognibene, C. Mathys, T. Fitzgerald, and G. Pezzulo (2015). "Active Inference and epistemic value." *Cognitive Neuroscience* 6(4): 187–214.

Friston, K., R. Rosch, T. Parr, C. Price, and H. Bowman (2017). "Deep temporal models and active inference." *Neuroscience and Biobehavioral Reviews* 77: 388–402.

Friston, K., S. Samothrakis and R. Montague (2012). "Active Inference and agency: optimal control without cost functions." *Biological Cybernetics* 106(8–9): 523–541.

Friston, K., P. Schwartenbeck, T. FitzGerald, M. Moutoussis, T. Behrens, and R. J. Dolan (2014). "The anatomy of choice: dopamine and decision-making." *Philosophical Transactions of the Royal Society B: Biological Sciences* 369(1655): 20130481.

Friston, K., K. Stephan, B. Li, and J. Daunizeau (2010). "Generalised filtering." *Mathematical Problems in Engineering*. doi:10.1155/2010/621670.

Friston, K., K. E. Stephan, R. Montague, and R. J. Dolan (2014). "Computational psychiatry: the brain as a phantastic organ." *Lancet Psychiatry* 1(2): 148–158.

Frith, C. D., S. Blakemore, and D. M. Wolpert (2000). "Explaining the symptoms of schizophrenia: abnormalities in the awareness of action." *Brain Research Reviews* 31(2–3): 357–363.

Funahashi, S., C. J. Bruce, and P. S. Goldman-Rakic (1989). "Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex." *Journal of Neurophysiology* 61(2): 331.

Fuster, J. n. M. (2004). "Upper processing stages of the perception-action cycle." *Trends in Cognitive Sciences* 8(4): 143–145.

Galea, J. M., S. Bestmann, M. Beigi, M. Jahanshahi, and J. C. Rothwell (2012). "Action reprogramming in Parkinson's disease: response to prediction error is modulated by levels of dopamine." *Journal of Neuroscience* 32(2): 542.

George, D., W. Lehrach, K. Kansky, M. Lázaro-Gredilla, C. Laan, B. Marthi, X. Lou, Z. Meng, Y. Liu, H. Wang, A. Lavin, and D. S. Phoenix (2017). "A generative vision model that trains with high data efficiency and breaks text-based CAPTCHAs." *Science* 358(6368): eaag2612.

Gershman, S. J., E. J. Horvitz, and J. B. Tenenbaum (2015). "Computational rationality: a converging paradigm for intelligence in brains, minds, and machines." *Science* 349(6245): 273.

Gertler, T. S., C. S. Chan, and D. J. Surmeier (2008). "Dichotomous anatomical properties of adult striatal medium spiny neurons." *Journal of Neuroscience* 28(43): 10814.

Gil, Z., B. W. Connors, and Y. Amitai (1997). "Differential regulation of neocortical synapses by neuromodulators and activity." *Neuron* 19(3): 679–686.

Goodfellow, I. J., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio (2014). "Generative adversarial networks." *arXiv:1406.2661 [cs, stat]*.

Gottlieb, J., P.-Y. Oudeyer, M. Lopes, and A. Baranes (2013). "Information-seeking, curiosity, and attention: computational and neural mechanisms." *Trends in Cognitive Sciences* 17(11): 585–593.

Gottwald, S., and D. A. Braun (2020). "The two kinds of free energy and the Bayesian revolution." *arXiv:2004.11763 [cs, q-bio]*.

Gregory, R. L. (1980). "Perceptions as hypotheses." *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 290(1038): 181–197.

Ha, D., and D. Eck (2017). "A neural representation of sketch drawings." *arXiv preprint arXiv:1704.03477*.

Ha, D., and J. Schmidhuber (2018). "World models." *arXiv:1803.10122 [cs, stat].*

Haeusler, S., and W. Maass (2007). "A statistical analysis of information-processing properties of lamina-specific cortical microcircuit models." *Cerebral Cortex* 17(1): 149–162.

Harlow, H. F. (1949). "The formation of learning sets." *Psychological Review* 56(1): 51–65.

Helmholtz, H. v. (1866). "Concerning the perceptions in general." *Treatise on Physiological Optics*. Translated by J. P. C. Southall. New York, Dover.

Helmholtz, H. v. (1867). *Handbuch der physiologischen Optik*. Leipzig: L. Voss.

Herbart, J. (1825). *Psychologie als Wissenschaft: Neu gegründet auf Erfahrung, Metaphysik und Mathematik*. Zweiter, analytischer Teil. Koenigsberg, Germany: August Wilhem Unzer.

Hezemans, F. H., N. Wolpe, and J. B. Rowe (2020). "Apathy is associated with reduced precision of prior beliefs about action outcomes." *Journal of Experimental Psychology: General* 149(9): 1767–1777.

Hills, T. T., P. M. Todd, D. Lazer, A. D. Redish, and I. D. Couzin (2015). "Exploration versus exploitation in space, mind, and society." *Trends in Cognitive Sciences* 19(1): 46–54.

Hillyard, S. A., E. K. Vogel, and S. J. Luck (1998). "Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence." *Philosophical Transactions of the Royal Society B: Biological Sciences* 353(1373): 1257–1270.

Hinton, G. E. (2007a). "Learning multiple layers of representation." *Trends in Cognitive Sciences* 11(10): 428–434.

Hinton, G. E. (2007b). "To recognize shapes, first learn to generate images." *Progress in Brain Research* 165: 535–547.

Hoffmann, J. (1993). *Vorhersage und Erkenntnis: Die Funktion von Antizipationen in der menschlichen Verhaltenssteuerung und Wahrnehmung* [Anticipation and cognition: The function of anticipations in human behavioral control and perception]. Goettingen, Germany: Hogrefe.

Hoffmann, J. (2003). "Anticipatory behavioral control." In *Anticipatory Behavior in Adaptive Learning Systems: Foundations, Theories, and Systems*, edited by M. V. Butz, O. Sigaud, and P. Gerard, 44–65. Berlin: Springer-Verlag.

Hohwy, J. (2013). *The Predictive Mind*. New York: Oxford University Press.

Hohwy, J. (2016). “The self-evidencing brain.” *Noûs* 50(2): 259–285.

Hommel, B., J. Musseler, G. Aschersleben, and W. Prinz (2001). “The theory of event coding (TEC): a framework for perception and action planning.” *Behavioral and Brain Science* 24(5): 849–878.

Huerta, R., and M. Rabinovich (2004). “Reproducible sequence generation in random neural ensembles.” *Physical Review Letters* 93(23): 238104.

Hurley, S. (2008). “The shared circuits model (SCM): how control, mirroring, and simulation can enable imitation, deliberation, and mindreading.” *Behavioral and Brain Sciences* 31: 1–22.

Huygens, C. (1673). *Horologium Oscillatorium: Sive, De Motu Pendulorum Ad Horologia Aptato Demostrationes Geometricae*. Culture et Civilisation.

Iodice, P., G. Porciello, I. Bufalari, L. Barca, and G. Pezzulo (2019). “An interoceptive illusion of effort induced by false heart-rate feedback.” *Proceedings of the National Academy of Sciences* 116(28): 13897–13902.

Isomura, T., and K. Friston (2018). “In vitro neural networks minimise variational free energy.” *Scientific Reports* 8(1): 16926.

Isomura, T., T. Parr, and K. Friston (2019). “Bayesian filtering with multiple internal models: toward a theory of social intelligence.” *Neural Computation* 31(12): 2390–2431.

James, W. (1890). *The Principles of Psychology*. New York: Dover Publications.

Jaynes, E. T. (1957). “Information theory and statistical mechanics.” *Physical Review* 106(4): 620.

Jeannerod, M. (2001). “Neural simulation of action: a unifying mechanism for motor cognition.” *NeuroImage* 14: S103–S109.

Joffily, M., and G. Coricelli (2013). “Emotional valence and the free-energy principle.” *PLOS Computational Biology* 9(6): e1003094.

Kahneman, D. (2017). Thinking, fast and slow. United Kingdom: Penguin Books.

Kakade, S., and P. Dayan (2002). “Dopamine: generalization and bonuses.” *Neural Networks* 15(4): 549–559.

Kanai, R., Y. Komura, S. Shipp, and K. Friston (2015). “Cerebral hierarchies: predictive processing, precision and the pulvinar.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 370(1668): 20140169.

Kaplan, R., and K. J. Friston (2018). “Planning and navigation as Active Inference.” *Biological Cybernetics* 112: 323–343.

Kappen, H. J., V. Gómez, and M. Opper (2012). “Optimal control as a graphical model inference problem.” *Machine Learning* 87(2): 159–182.

Karson, C. N. (1983). “Spontaneous eye-blink rates and dopaminergic systems.” *Brain* 106(3): 643–653.

Kemp, C., and J. B. Tenenbaum (2008). “The discovery of structural form.” *Proceedings of the National Academy of Sciences* 105(31): 10687–10692.

Kiebel, S. J., J. Daunizeau, and K. J. Friston (2008). “A hierarchy of time-scales and the brain.” *PLOS Computational Biology* 4(11): e1000209.

Kingma, D. P., and M. Welling (2014). “Auto-encoding variational Bayes.” *arXiv:1312.6114 [cs, stat].*

Kirchhoff, M., T. Parr, E. Palacios, K. Friston, and J. Kiverstein (2018). “The Markov blankets of life: autonomy, active inference and the free energy principle.” *Journal of the Royal Society*, Interface 15, 20170792, doi:10.1098/rsif.2017.0792.

Knill, D. C., and A. Pouget (2004). “The Bayesian brain: The role of uncertainty in neural coding and computation.” *Trends in Neurosciences* 27(12): 712–719.

Kording, K. P., and D. M. Wolpert (2006). “Bayesian decision theory in sensorimotor control.” *Trends in Cognitive Sciences* 10: 319–326.

Koss, M. C. (1986). “Pupillary dilation as an index of central nervous system α2-adrenoceptor activation.” *Journal of Pharmacological Methods* 15(1): 1–19.

Krakauer, J. W., A. A. Ghazanfar, A. Gomez-Marin, M. A. MacIver, and D. Poeppel (2017). “Neuroscience needs behavior: correcting a reductionist bias.” *Neuron* 93(3): 480–490.

Krishnamurthy, K., M. R. Nassar, S. Sarode, and J. I. Gold (2017). “Arousal-related adjustments of perceptual biases optimize perception in dynamic environments.” *Nature Human Behaviour* 1: 0107.

Kunde, W., I. Koch, and J. Hoffmann (2004). “Anticipated action effects affect the selection, initiation and execution of actions.” *Quarterly Journal of Experimental Psychology. Section A: Human Experimental Psychology* 57(1): 87–106.

Lake, B. M., T. D. Ullman, J. B. Tenenbaum, and S. J. Gershman (2017). “Building machines that learn and think like people.” *Behavioral and Brain Sciences* 40: 1–72.

Lambe, E. K., P. S. Goldman-Rakic, and G. K. Aghajanian (2000). “Serotonin induces EPSCs preferentially in layer V pyramidal neurons of the frontal cortex in the rat.” *Cerebral Cortex* 10(10): 974–980.

Lavín, C., R. San Martín, and E. Rosales Jubal (2013). “Pupil dilation signals uncertainty and surprise in a learning gambling task.” *Frontiers in Behavioral Neuroscience* 7: 218.

Lavine, N., M. Reuben, and P. Clarke (1997). “A population of nicotinic receptors is associated with thalamocortical afferents in the adult rat: laminal and areal analysis.” *Journal of Comparative Neurology* 380(2): 175–190.

Lee, M. D., and E.-J. Wagenmakers (2014). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge: Cambridge University Press.

Lee, S. W., S. Shimojo, and J. P. O'Doherty (2014). "Neural computations underlying arbitration between model-based and model-free learning." *Neuron* 81(3): 687–699.

Levine, S. (2018). "Reinforcement learning and control as probabilistic inference: tutorial and review." *arXiv:1805.00909 [cs, stat].*

Liao, H.-I., M. Yoneya, S. Kidani, M. Kashino, and S. Furukawa (2016). "Human pupillary dilation response to deviant auditory stimuli: effects of stimulus properties and voluntary attention." *Frontiers in Neuroscience* 10: 43.

Limanowski, J., and K. Friston (2019). "Attentional modulation of vision versus proprioception during action." *Cerebral Cortex* 30(3): 1637–1648.

Lindley, D. V. (1956). "On a measure of the information provided by an experiment." *Annals of Mathematical Statistics* 27(4): 986–1005.

Linson, A., T. Parr, and K. J. Friston (2020). "Active Inference, stressors, and psychological trauma: a neuroethological model of (mal)adaptive explore-exploit dynamics in ecological context." *Behavioural Brain Research* 380: 1–13.

Loeliger, H. A. (2004). "An introduction to factor graphs." *IEEE Signal Processing Magazine* 21(1): 28–41.

Loeliger, H. A., J. Dauwels, J. Hu, S. Korl, L. Ping, and F. R. Kschischang (2007). "The factor graph approach to model-based signal processing." *Proceedings of the IEEE* 95(6): 1295–1322.

MacKay, D. M. (1956). *The Epistemological Problem for Automata*. Princeton, NJ: Princeton University Press.

Maisto, D., L. Barca, O. V. d. Bergh, and G. Pezzulo (2021). "Perception and misperception of bodily symptoms from an Active Inference perspective: modelling the case of panic disorder." *Psychological Review*.

Maisto, D., K. Friston, and G. Pezzulo (2019). "Caching mechanisms for habit formation in Active Inference." *Neurocomputing* 359: 298–314.

Marek, R., C. Strobel, T. W. Bredy, and P. Sah (2013). "The amygdala and medial prefrontal cortex: partners in the fear circuit." *Journal of Physiology* 591(10): 2381–2391.

Marshall, L., C. Mathys, D. Ruge, A. O. de Berker, P. Dayan, K. E. Stephan, and S. Bestmann (2016). "Pharmacological fingerprints of contextual uncertainty." *PLOS Biology* 14(11): e1002575.

Maturana, H. R., and F. J. Varela (1980). *Autopoiesis and Cognition: The Realization of Living*. Dordrecht, Holland: D. Reidel.

Mesulam, M. M. (1998). "From sensation to cognition." *Brain: Journal of Neurology* 121(pt. 6): 1013–1052.

Miller, E. K., and J. D. Cohen (2001). "An integrative theory of prefrontal cortex function." *Annual Review of Neuroscience* 24: 167–202.

Miller, G. A., E. Galanter, and K. H. Pribram (1960). *Plans and the Structure of Behavior*. New York: Holt, Rinehart and Winston.

Miller, K. D. (2003). "Understanding layer 4 of the cortical circuit: a model based on cat V1." *Cerebral Cortex* 13(1): 73–82.

Millidge, B. (2019). "Deep Active Inference as variational policy gradients." *arXiv:1907.03876 [cs]*.

Mirza, M. B., R. A. Adams, C. Mathys, and K. J. Friston (2018). "Human visual exploration reduces uncertainty about the sensed world." *PLOS ONE* 13(1): e0190429.

Mirza, M. B., R. A. Adams, C. D. Mathys, and K. J. Friston (2016). "Scene construction, visual foraging, and Active Inference." *Frontiers in Computational Neuroscience* 10(56).

Mirza, M. B., R. A. Adams, T. Parr, and K. Friston (2019). "Impulsivity and Active Inference." *Journal of Cognitive Neuroscience* 31(2): 202–220.

Montague, P. R., R. J. Dolan, K. J. Friston, and P. Dayan (2012). "Computational psychiatry." *Trends in Cognitive Sciences* 16(1): 72–80.

Moran, R. J., P. Campo, M. Symmonds, K. E. Stephan, R. J. Dolan, and K. J. Friston (2013). "Free energy, precision and learning: the role of cholinergic neuromodulation." *Journal of Neuroscience* 33(19): 8227–8236.

Moss, J., and J. P. Bolam (2008). "A dopaminergic axon lattice in the striatum and its relationship with cortical and thalamic terminals." *Journal of Neuroscience* 28(44): 11221.

Moutoussis, M., N. J. Trujillo-Barreto, W. El-Deredy, R. J. Dolan, and K. J. Friston (2014). "A formal model of interpersonal inference." *Frontiers in Human Neuroscience* 8: 160.

Mukherjee, P., A. Sabharwal, R. Kotov, A. Szekely, R. Parsey, D. M. Barch, and A. Mohanty (2016). "Disconnection between amygdala and medial prefrontal cortex in psychotic disorders." *Schizophrenia Bulletin* 42(4): 1056–1067.

Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. Cambridge, MA: MIT Press.

Nambu, A. (2004). "A new dynamic model of the cortico-basal ganglia loop." *Progress in Brain Research* 143: 461–466.

Nassar, M. R., K. M. Rumsey, R. C. Wilson, K. Parikh, B. Heasly, and J. I. Gold (2012). "Rational regulation of learning dynamics by pupil-linked arousal systems." *Nature Neuroscience* 15(7): 1040–1046.

Nave, K., G. Deane, M. Miller, and A. Clark (2020). "Wilding the predictive brain." *Cognitive Science* 11(6): e1542.

Neisser, U. (2014). *Cognitive Psychology: Classic Edition*. London: Taylor & Francis.

Nishimoto, R., and J. Tani (2009). "Development of hierarchical structures for actions and motor imagery: a constructivist view from synthetic neuro-robotics study." *Psychological Research PRPF* 73(4): 545–558.

Olsen, S. R., D. S. Bortone, H. Adesnik, and M. Scanziani (2012). "Gain control by layer six in cortical circuits of vision." *Nature* 483: 47.

Ortega, P. A., and D. A. Braun (2013). "Thermodynamics as a theory of decision-making with information-processing costs." *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* 469(2153).

Oudeyer, P. Y., F. Kaplan, and V. Hafner (2007). "Intrinsic motivation systems for autonomous mental development." *IEEE Transactions on Evolutionary Computation* 11(2): 265–286.

Palacios, E. R., T. Isomura, T. Parr, and K. Friston (2019). "The emergence of synchrony in networks of mutually inferring neurons." *Scientific Reports* 9(1): 6412.

Palacios, E. R., A. Razi, T. Parr, M. Kirchhoff, and K. Friston (2020). "On Markov blankets and hierarchical self-organisation." *Journal of Theoretical Biology* 486: 110089.

Pareés, I., H. Brown, A. Nuruki, R. A. Adams, M. Davare, K. P. Bhatia, K. Friston, and M. J. Edwards (2014). "Loss of sensory attenuation in patients with functional (psychogenic) movement disorders." *Brain* 137(11): 2916–2921.

Parr, T. (2020). "Inferring what to do (and what not to)." *Entropy* 22(5): 536.

Parr, T., D. A. Benrimoh, P. Vincent, and K. J. Friston (2018). "Precision and false perceptual inference." *Frontiers in Integrative Neuroscience* 12: 39–39.

Parr, T., L. D. Costa, and K. Friston (2020). "Markov blankets, information geometry and stochastic thermodynamics." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 378(2164): 20190159.

Parr, T., and K. J. Friston (2017a). "The computational anatomy of visual neglect." *Cerebral Cortex* 28: 1–14.

Parr, T., and K. J. Friston (2017b). "Uncertainty, epistemics and active inference." *Journal of the Royal Society Interface* 14(136).

Parr, T., and K. J. Friston (2017c). "Working memory, attention, and salience in Active Inference." *Scientific Reports* 7(1): 14678.

Parr, T., and K. J. Friston (2018a). “Active Inference and the anatomy of oculomotion.” *Neuropsychologia* 111: 334–343.

Parr, T., and K. J. Friston (2018b). “The anatomy of inference: generative models and brain structure.” *Frontiers in Computational Neuroscience* 12(90).

Parr, T., and K. J. Friston (2018c). “The discrete and continuous brain: From decisions to movement—and back again.” *Neural Computation* 30(9): 2319–2347.

Parr, T., and K. J. Friston (2018d). “Generalised free energy and Active Inference: can the future cause the past?” *bioRxiv*.

Parr, T., and K. J. Friston (2019a). “Attention or salience?” *Current Opinion in Psychology* 29: 1–5.

Parr, T., and K. J. Friston (2019b). “The computational pharmacology of oculomotion.” *Psychopharmacology* 236(8): 2473–2484.

Parr, T., D. Markovic, S. J. Kiebel, and K. J. Friston (2019). “Neuronal message passing using mean-field, Bethe, and marginal approximations.” *Scientific Reports* 9(1): 1889.

Parr, T., M. B. Mirza, H. Cagnan, and K. J. Friston (2019). “Dynamic causal modelling of active vision.” *Journal of Neuroscience* 39(32): 6265–6275.

Parr, T., R. V. Rikhye, M. M. Halassa, and K. J. Friston (2019). “Prefrontal computation as Active Inference.” *Cerebral Cortex* 30(2): 682–695.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, CA: Morgan Kaufmann.

Pearl, J. and D. Mackenzie (2018). *The Book of Why: The New Science of Cause and Effect*. New York: Basic Books.

Perrinet, L. U., R. A. Adams, and K. J. Friston (2014). “Active Inference, eye movements and oculomotor delays.” *Biological Cybernetics* 108(6): 777–801.

Peters, A., B. S. McEwen, and K. Friston (2017). “Uncertainty and stress: why it causes diseases and how it is mastered by the brain.” *Progress in Neurobiology*. 156: 164–188

Petersen, K. B., and M. S. Pedersen (2012). *The Matrix Cookbook*. https://www.math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf.

Pezzulo, G. (2012). “An Active Inference view of cognitive control.” *Frontiers in Theoretical and Philosophical Psychology* 478: 1–2

Pezzulo, G. (2013). “Why do you fear the bogeyman? An embodied predictive coding model of perceptual inference.” *Cognitive, Affective, and Behavioral Neuroscience* 14(3): 902–911.

Pezzulo, G., G. Baldassarre, M. V. Butz, C. Castelfranchi, and J. Hoffmann (2007). “Fron action to goals and vice-versa: Theoretical analysis and models of the ideomotor principle and TOTE." In *Anticipatory Behavior in Adaptive Learning Systems*, edited by M. V. Butz, O. Sigaud, G. Pezzulo, and G. Baldassarre. ABiALS 2006. *Lecture Notes in Computer Science*, vol 4520. Berlin: Springer. https://doi.org/10.1007/978-3-540-74262-3_5.

Pezzulo, G., L. W. Barsalou, A. Cangelosi, M. H. Fischer, K. McRae, and M. J. Spivey (2013). "Computational grounded cognition: a new alliance between grounded cognition and computational modeling." *Frontiers in Psychology* 3: 612.

Pezzulo, G., E. Cartoni, F. Rigoli, L. Pio-Lopez, and K. Friston (2016). "Active Inference, epistemic value, and vicarious trial and error." *Learning and Memory* 23(7): 322–338.

Pezzulo, G., and P. Cisek (2016). "Navigating the affordance landscape: feedback control as a process model of behavior and cognition." *Trends in Cognitive Sciences* 20(6): 414–424.

Pezzulo, G., F. Donnarumma, P. Iodice, D. Maisto, and I. Stoianov (2017). "Model-based approaches to active perception and control." *Entropy* 19(6): 266.

Pezzulo, G., C. Kemere, and M. A. A. van der Meer (2017). "Internally generated hippocampal sequences as a vantage point to probe future-oriented cognition." *Annals of the New York Academy of Sciences* 1396(1): 144–165.

Pezzulo, G., and M. Levin (2015). "Re-membering the body: applications of computational neuroscience to the top-down control of regeneration of limbs and other complex organs." *Integrative Biology* 7(12): 1487–1517.

Pezzulo, G., B. Lw, A. Cangelosi, M. H. Fischer, K. McRae, and M. Spivey (2011). "The mechanics of embodiment: A dialogue on embodiment and computational modeling." *Frontiers in Cognition* 2(5): 1–21.

Pezzulo, G., D. Maisto, L. Barca, and O. V. d. Bergh (2019). "Symptom perception from a predictive processing perspective." *Clinical Psychology in Europe* 1(4): 1–14.

Pezzulo, G., and F. Rigoli (2011). "The value of foresight: how prospection affects decision-making." *Frontiers in Neuroscience* 5(79).

Pezzulo, G., F. Rigoli, and K. J. Friston (2015). "Active Inference, homeostatic regulation and adaptive behavioural control." *Progress in Neurobiology* 136: 17–35.

Pezzulo, G., F. Rigoli, and K. J. Friston (2018). "Hierarchical Active Inference: a theory of motivated control." *Trends in Cognitive Sciences* 22(4): 294–306.

Pezzulo, G., M. Zorzi, and M. Corbetta (2020). "The secret life of predictive brains: what's spontaneous activity for?" *psyarxiv*.

Pfeifer, R., and J. C. Bongard (2006). *How the Body Shapes the Way We Think*. Cambridge, MA: MIT Press.

Pio-Lopez, L., A. Nizard, K. Friston, and G. Pezzulo (2016). "Active Inference and robot control: a case study." *Journal of the Royal Society Interface* 13(122).

Posner, M. I., R. D. Rafal, L. S. Choate, and J. Vaughan (1985). "Inhibition of return: neural basis and function." *Cognitive Neuropsychology* 2(3): 211–228.

Pouget, A., J. M. Beck, W. J. Ma, and P. E. Latham (2013). "Probabilistic brains: knowns and unknowns." *Nature Neuroscience* 16(9): 1170–1178.

Powers, W. T. (1973). *Behavior: The Control of Perception*. Hawthorne, NY: Aldine.

Prosser, A., K. J. Friston, N. Bakker, and T. Parr (2018). "A Bayesian account of psychopathy: a model of lacks remorse and self-aggrandizing." *Computational Psychiatry* 1–49.

Ramstead, M. J. D., M. D. Kirchhoff, and K. J. Friston (2019). "A tale of two densities: Active Inference is enactive inference." *Adaptive Behavior* 28(4): 225–239.

Rao, R. P., and D. H. Ballard (1999). "Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects." *Nature Neuroscience* 2(1): 79–87.

Rawlik, K., M. Toussaint, and S. Vijayakumar (2013). "On stochastic optimal control and reinforcement learning by approximate inference." In *Robotics: Science and Systems VIII*, edited by N. Roy, P. Newman, and S. Srinivasa. Cambridge, MA: MIT Press.

Risken, H. (1996). "Fokker-Planck equation." *The Fokker-Planck Equation: Methods of Solution and Applications*, 63–95. Berlin: Springer.

Rizzolatti, G., L. Riggio, I. Dascola, and C. Umiltá (1987). "Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention." *Neuropsychologia* 25(1, pt. 1): 31–40.

Rosenblueth, A., N. Wiener, and J. Bigelow (1943). "Behavior, purpose and teleology." *Philosophy of Science* 10(1): 18–24.

Sahin, M., W. D. Bowen, and J. P. Donoghue (1992). "Location of nicotinic and muscarinic cholinergic and μ-opiate receptors in rat cerebral neocortex: evidence from thalamic and cortical lesions." *Brain Research* 579(1): 135–147.

Sales, A. C., K. J. Friston, M. W. Jones, A. E. Pickering, and R. J. Moran (2019). "Locus coeruleus tracking of prediction errors optimises cognitive flexibility: an Active Inference model." *PLOS Computational Biology* 15(1): e1006267.

Sancaktar, C., M. van Gerven, and P. Lanillos (2020). "End-to-end pixel-based deep Active Inference for body perception and action." *arXiv:2001.05847 [cs, q-bio]*.

Schmidhuber, J. (1991). "Adaptive confidence and adaptive curiosity." Institut fur Informatik, Technische Universitat Munchen.

Schultz, W., P. Dayan, and P. R. Montague (1997). "A neural substrate of prediction and reward." *Science* 275(5306): 1593.

Schwartenbeck, P., T. H. B. FitzGerald, C. Mathys, R. Dolan, and K. Friston (2015). "The dopaminergic midbrain encodes the expected certainty about desired outcomes." *Cerebral Cortex* 25(10): 3434–3445.

Schwartenbeck, P., T. H. B. FitzGerald, C. Mathys, R. Dolan, F. Wurst, M. Kronbichler, and K. Friston (2015). "Optimal inference with suboptimal models: Addiction and active Bayesian inference." *Medical Hypotheses* 84(2): 109–117.

Schwartenbeck, P., and K. Friston (2016). "Computational phenotyping in psychiatry: a worked example." *eNeuro* 3(4): eneuro.0049–0016.2016.

Schwartenbeck, P., J. Passecker, T. U. Hauser, T. H. FitzGerald, M. Kronbichler, and K. J. Friston (2019). "Computational mechanisms of curiosity and goal-directed exploration." *eLife* 8: e41703.

Schwöbel, S., S. Kiebel, and D. Marković (2018). "Active Inference, belief propagation, and the Bethe approximation." *Neural Computation* 30(9): 1–38.

Seth, A. K. (2013). "Interoceptive inference, emotion, and the embodied self." *Trends in Cognitive Sciences* 17(11): 565–573.

Seth, A. K., and K. J. Friston (2016). "Active interoceptive inference and the emotional brain." *Philosophical Transactions of the Royal Society B* 371(1708): 20160007.

Seth, A. K., K. Suzuki, and H. D. Critchley (2012). "An interoceptive predictive coding model of conscious presence." *Frontiers in Psychology* 2: 1–16.

Shadmehr, R., M. A. Smith, and J. W. Krakauer (2010). "Error correction, sensory prediction, and adaptation in motor control." *Annual Review of Neuroscience* 33: 89–108.

Sheliga, B. M., L. Riggio, and G. Rizzolatti (1994). "Orienting of attention and eye movements." *Experimental Brain Research* 98(3): 507–522.

Sheliga, B. M., L. Riggio, and G. Rizzolatti (1995). "Spatial attention and eye movements." *Experimental Brain Research* 105(2): 261–275.

Shipp, S. (2007). "Structure and function of the cerebral cortex." *Current Biology* 17(12): R443–R449.

Shipp, S. (2016). "Neural elements for predictive coding." *Frontiers in Psychology* 7: 1792.

Shipp, S., R. A. Adams, and K. J. Friston (2013). "Reflections on agranular architecture: predictive coding in the motor cortex." *Trends in Neurosciences* 36(12): 706–716.

Simon, H. A. (1990). "Bounded rationality." In *Utility and Probability*, 15–18. New York: Springer.

Skinner, B. F. (1938). *The Behavior of Organisms: An Experimental Analysis*. New York: Appleton-Century-Crofts.

Smith, R., R. D. Lane, T. Parr, and K. J. Friston (2019). "Neurocomputational mechanisms underlying emotional awareness: insights afforded by deep Active Inference and their potential clinical relevance." *Neuroscience and Biobehavioral Reviews* 107: 473–491.

Smith, R., T. Parr, and K. J. Friston (2019). "Simulating emotions: an Active Inference model of emotional state inference and emotion concept learning." *bioRxiv* 640813.

Solway, A., and M. M. Botvinick (2012). "Goal-directed decision making as probabilistic inference: a computational framework and potential neural correlates." *Psychological Review* 119(1): 120–154.

Sterling, P. (2012). "Allostasis: a model of predictive regulation." *Physiology and Behavior* 106(1): 5–15.

Stewart, N., N. Chater, and G. D. A. Brown (2006). "Decision by sampling." *Cognitive Psychology* 53(1): 1–26.

Stoianov, I., D. Maisto, and G. Pezzulo (2020). "The hippocampal formation as a hierarchical generative model supporting generative replay and continual learning." *bioRxiv* 2020.2001.2016.908889.

Sutton, R. S., and A. G. Barto (1998). *Reinforcement Learning: An Introduction*. Cambridge MA: MIT Press.

Tani, J., and J. White (2020). "Cognitive neurorobotics and self in the shared world, a focused review of ongoing research." *Adaptive Behavior*. 1–20.

Tenenbaum, J. B., T. L. Griffiths, and C. Kemp (2006). "Theory-based Bayesian models of inductive learning and reasoning." *Trends in Cognitive Sciences* 10: 309–318.

Tervo, D. G. R., J. B. Tenenbaum, and S. J. Gershman (2016). "Toward the neural implementation of structure learning." *Current Opinion in Neurobiology* 37: 99–105.

Thomson, A. (2010). "Neocortical layer 6, a review." *Frontiers in Neuroanatomy* 4(13).

Todorov, E. (2004). "Optimality principles in sensorimotor control." *Nature Neuroscience* 7(9): 907–915.

Todorov, E. (2008). "General duality between optimal control and estimation." In *47th IEEE Conference on Decision and Control*, 4286–4292.

Todorov, E. (2009). "Efficient computation of optimal actions." *Proceedings of the National Academy of Sciences USA* 106(28): 11478–11483.

Tolman, E. C. (1948). "Cognitive maps in rats and men." *Psychological Review* 55: 189–208.

Tschantz, A., L. Barca, D. Maisto, C. L. Buckley, A. K. Seth, and G. Pezzulo (2021). "Simulating homeostatic, allostatic and goal-directed forms of interoceptive control using Active Inference." *bioRxiv* 2021.2002.2016.431365.

Tschantz, A., A. K. Seth, and C. L. Buckley (2020). "Learning action-oriented models through active inference." *PLOS Computational Biology* 16(4): e1007805.

Tsvetanov, K. A., R. N. A. Henson, L. K. Tyler, A. Razi, L. Geerligs, T. E. Ham, and J. B. Rowe (2016). "Extrinsic and intrinsic brain network connectivity maintains cognition across the lifespan despite accelerated decay of regional brain activation." *Journal of Neuroscience* 36(11): 3115.

Ueltzhöffer, K. (2018). "Deep Active Inference." *Biological Cybernetics* 112(6): 547–573.

Ungerleider, L. G., and J. V. Haxby (1994). "'What' and 'where' in the human brain." *Current Opinion in Neurobiology* 4(2): 157–165.

van de Laar, T. W., and B. de Vries (2019). "Simulating Active Inference processes by message passing." *Frontiers in Robotics and AI* 6(20).

Veissière, S. P. L., A. Constant, M. J. D. Ramstead, K. J. Friston, and L. J. Kirmayer (2020). "Thinking through other minds: a variational approach to cognition and culture." *Behavioral and Brain Sciences* 43: e90.

Verschure, P., C. M. A. Pennartz, and G. Pezzulo (2014). The why, what, where, when and how of goal-directed choice: neuronal and computational principles. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 369: 20130483.

Verschure, P. F. M. J. (2012). "Distributed adaptive control: a theory of the mind, brain, body nexus." *Biologically Inspired Cognitive Architectures* 1: 55–72.

Verschure, P. F. M. J., T. Voegtlin, and R. J. Douglas (2003). "Environmentally mediated synergy between perception and behaviour in mobile robots." *Nature* 425(6958): 620–624.

Vincent, P., T. Parr, D. Benrimoh, and K. J. Friston (2019). "With an eye on uncertainty: modelling pupillary responses to environmental volatility." *PLOS Computational Biology* 15(7): e1007126.

Vossel, S., M. Bauer, C. Mathys, R. A. Adams, R. J. Dolan, K. E. Stephan, and K. J. Friston (2014). "Cholinergic stimulation enhances Bayesian belief updating in the deployment of spatial attention." *Journal of Neuroscience* 34(47): 15735.

Wainwright, M. J., and M. I. Jordan (2008). "Graphical models, exponential families, and variational inference." *Foundations and Trends in Machine Learning* 1(1–2): 1–305.

Wald, A. (1947). "An essentially complete class of admissible decision functions." *Annals of Mathematical Statistics* 18(4): 549–555.

Wall, N. R., M. De La Parra, E. M. Callaway, and A. C. Kreitzer (2013). "Differential innervation of direct- and indirect-pathway striatal projection neurons." *Neuron* 79(2): 347–360.

Wesson, D. W., and D. A. Wilson (2011). "Sniffing out the contributions of the olfactory tubercle to the sense of smell: hedonics, sensory integration, and more?" *Neuroscience and Biobehavioral Reviews* 35(3): 655–668.

Wiener, N. (1948). *Cybernetics: or Control and Communication in the Animal and the Machine*. Cambridge, MA: MIT Press.

Winn, J., and C. M. Bishop (2005). "Variational message passing." *Journal of Machine Learning Research* 6(April): 661–694.

Wolpert, D. M., K. Doya, and M. Kawato (2003). "A unifying computational framework for motor control and social interaction." *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 358(1431): 593–602.

Wolpert, D. M., and M. Kawato (1998). "Multiple paired forward and inverse models for motor control." *Neural Networks* 11(7–8): 1317–1329.

Wolpert, D. M., and M. S. Landy (2012). "Motor control is decision-making." *Current Opinion in Neurobiology* 22(6): 996–1003.

Yager, L. M., A. F. Garcia, A. M. Wunsch, and S. M. Ferguson (2015). "The ins and outs of the striatum: role in drug addiction." *Neuroscience* 301: 529–541.

Yamashita, Y., and J. Tani (2008). "Emergence of functional hierarchy in a multiple timescale neural network model: a humanoid robot experiment." *PLOS Computational Biology* 4(11): e1000220.

Yuan, R., and P. Ao (2012). "Beyond Itô versus Stratonovich." *Journal of Statistical Mechanics: Theory and Experiment* 2012(07): P07010.

Yuille, A., and D. Kersten (2006). "Vision as Bayesian inference: analysis by synthesis?" *Probabilistic Models of Cognition* 10(7): 301–308.

Zeki, S., and S. Shipp (1988). "The functional logic of cortical connections." *Nature* 335(6188): 311–317.

Zénon, A., O. Solopchuk, and G. Pezzulo (2019). "An information-theoretic perspective on the costs of cognition." *Neuropsychologia* 123: 5–18.

Zhang, Z., S. Cordeiro Matos, S. Jego, A. Adamantidis, and P. Séguéla (2013). "Norepinephrine drives persistent activity in prefrontal cortex via synergistic α_1 and α_2 adrenoceptors." *PLOS ONE* 8(6): e66122.

Zhou, Y., P. Zeidman, S. Wu, A. Razi, C. Chen, L. Yang, J. Zou, G. Wang, H. Wang, and K. J. Friston (2018). "Altered intrinsic and extrinsic connectivity in schizophrenia." *NeuroImage: Clinical* 17: 704–716.