Emotion and Decision-Making Explained

Edmund T. Rolls
Oxford Centre for Computational Neuroscience
Oxford
England

OXFORD UNIVERSITY PRESS • OXFORD 2014

Preface

What produces emotions? Why do we have emotions? How do we have emotions? Why do emotional states feel like something? How do we take decisions? This book seeks explanations of emotion and decision-making by considering these questions.

One of the distinctive properties of this book is that it develops a conceptual and evolutionary approach to emotion (see for example Chapters 2 and 3). This approach shows how cognitive states can produce and modulate emotion, and in turn how emotional states can influence cognition. Another distinctive property is that this book links these approaches to studies on the brain, at the level of neuronal neurophysiology, which provides much of the primary data about how the brain operates; but also to neuropsychological studies of patients with brain damage; to functional magnetic resonance imaging (fMRI) (and other neuroimaging) approaches; and to computational neuroscience approaches. The author's research is in all these areas, and this may help the approach to emotion and decision-making described here to span many levels of investigation. Another distinctive property of this book is that it extends the search beyond emotional value, to economic value, to how decisions are then taken, between for example stimuli that have different value. The empirical evidence that is brought to bear is largely from non-human primates and from humans, because of the considerable similarity of their visual and emotional systems associated with the great development of the prefrontal cortex and temporal lobes in primates, and because the overall aim is to understand how emotion and decision-making are implemented in the human brain, and the neuropsychiatric disorders that can arise.

To understand how the brain works, including how it functions in emotion and decision-making, it is necessary to combine different approaches, including neural computation. Neurophysiology at the single neuron level is needed because this is the level at which information is exchanged between the computing elements of the brain. Evidence from the effects of brain damage, including that available from neuropsychology, is needed to help understand what different parts of the system do, and indeed what each part is necessary for. Neuroimaging is useful to indicate where in the human brain different processes take place, and to show which functions can be dissociated from each other. Knowledge of the biophysical and synaptic properties of neurons is essential to understand how the computing elements of the brain work, and therefore what the building blocks of biologically realistic computational models should be. Knowledge of the anatomical and functional architecture of the cortex is needed to show what types of neuronal network actually perform the computation. And finally the approach of neural computation is needed, as this is required to link together all the empirical evidence to produce an understanding of how the system actually works. This book utilizes evidence from all these disciplines to develop an understanding of how emotion and decision-making are implemented by processing in the brain.

The overall plan of the book is as follows. Chapter 1 outlines the ways in which this book approaches different types of explanation of emotion, and introduces some of the concepts. Chapter 2 then considers the nature of emotion, producing a theory of emotion, and comparing it to some other theories. Chapter 3 considers the functions of emotion, and leads to a Darwinian theory of the adaptive value of emotion, which helps to illuminate many aspects of brain design and behaviour. Chapter 4 takes the explanation of emotion to the level of how emotion is implemented in the brain. Chapter 5 extends and complements this by

extending the approach to motivated behaviour, in which affective responses to sensory stimuli are fundamental in for example appetite and the control of food intake. Chapter 6 extends the approach to the pharmacology of emotion and addiction. Chapter 7 extends the approach further, to sexual behaviour. Chapter 8 then considers how one proceeds beyond assessing affective value, by describing brain mechanisms involved in taking decisions between stimuli of for example different affective value, or between different sensory stimuli. It turns out that the mechanisms of decision-making are used in many different brain areas for different purposes, including perceptual categorization, decisions about actions, memory recall, and short-term memory, and the mechanisms described encompass all of these in a unifying conceptual approach. Chapter 8 also considers confidence in decisions, and how we may be able to correct a decision even before the outcome of the choice has been provided. Chapter 9 describes factors that influence decision-making between stimuli or 'goods' of different value, by describing findings in the field of neuroeconomics. Chapter 10 then considers the issue of emotional feelings, which is part of the much larger issue of consciousness. Chapter 11 then synthesizes some of the points made, including how decisions are made and are influenced by emotions.

Appendix A describes some of the computational framework for understanding how systems in the brain in the form of neural networks perform emotion-related learning. Appendix B provides an overview of many different approaches to decision-making, comparing phenomenological with mechanistic approaches. The treatment in Appendix B includes quantitative descriptions of many of these approaches, including the biologically plausible integrate-and-fire attractor model of decision-making (Section B.5) and its mean-field equivalent (Section B.6). Appendix C provides a Glossary of some of the terms used in the book. The book thus seeks to explain emotions in terms of the following: What produces emotions? Why do we have emotions? How do we have emotions? Why do emotional states feel like something? How do we take decisions?

This book evolved from my earlier books *The Brain and Emotion* (Rolls 1999a) and *Emotion Explained* (Rolls 2005b) in some of the following ways:

Emotion and Decision-Making Explained (2014) updates *Emotion Explained* (2005) with much recent research on emotion and reward value, and then goes beyond this by a treatment in Chapter 9 of economic value from the field of neuroeconomics, and then goes further to provide a substantial treatment of decision-making (Chapter 8 and Appendix B), which includes decision-making between stimuli or goods of different value, but also applies to many types of decision-making. This is a natural extension of my treatment of emotion, for after value has been assessed, choices must be made, both between reinforcing stimuli, and about actions to take given the benefits and the costs.

Emotion and Decision-Making Explained goes beyond the brain mechanisms of emotion, in that it seeks to explain emotions in terms of the following: What produces emotions? (The general answer I propose is rewards and punishers, but with other factors too.) Why do we have emotions? (The overall answer I propose is that emotions are evolutionarily adaptive as they provide an efficient way for genes to influence our behaviour to increase their success.) How do we have emotions? (I answer this by describing what is known about the brain mechanisms of emotion.) Why do emotional states feel like something? This is part of the large problem of consciousness, which I address in Chapter 10.

Emotion and Decision-Making Explained also goes beyond the brain mechanisms of emotion by developing my approach and theory of the nature of emotion, and comparing my approach to a range of different approaches to the nature of emotion, including the approaches of Antonio Damasio, Joseph LeDoux, Jaak Panksepp, and appraisal theorists such as Klaus Scherer.

Another way in which this book goes beyond brain mechanisms of emotion is to propose in Chapter 3 a Darwinian account of why animals (including humans) have emotions. The theory will I believe stand the test of time, in the same way as Darwin's theory of evolution by natural selection, and argues that emotions have the important evolutionary role of enabling genes to specify the goals (i.e. the rewards etc. that produce emotions) for actions, rather than the actions themselves. The advantage of this Darwinian design is that although the genes specify the goals, the actual actions are not prespecified by the genes, so that there is great flexibility of the actions themselves. This provides a new approach to the nature vs nurture debate in animal behaviour, for it shows how genes can influence behaviour without specifying a fixed, instinctive, behavioural response. I hope that this will make the book of interest to a wide audience, including many interested in evolution and evolutionary biology.

Although in evolution Darwinian processes lead to gene-defined goals, it is also the case that in humans, goals may be influenced by other processes, including cultural processes. Indeed, some goals are defined within a culture, for example writing a novel like one by Tolstoy vs one by Virginia Woolf. But it is argued that it is primary reinforcers specified by genes of the general type shown in Table 2.1 on page 20 that make us want to be recognized in society because of the advantages this can bring, to solve difficult problems, etc., and therefore to perform actions such as writing novels (see further Ridley (2003) Chapter 8, Ridley (1993b) pp. 310 ff, Laland & Brown (2002) pp. 271 ff, and Dawkins (1982)). Indeed, culture is influenced by human genetic propensities, and it follows that human cognitive, affective, and moral capacities are the product of a unique dynamic known as *gene-culture coevolution* (Gintis 2011, Gintis 2007).

We may also note that the theory that genes set many goals for action does not mean that our behaviour is determined by genes. Modern evolutionary theory has led to the understanding that many traits, particularly behavioural ones, may have some genetic basis but that does not mean that they will inevitably appear, because much depends on the environment (Dawkins 1995, Ridley 2003). Further, part of the power of the theory of emotion described here is that in evolution genes specify rewards and punishers that are goals for action, but do not specify the actions themselves, which are flexible and can be learned. Further, it is shown in Chapter 10 that in humans (and other animals) with a reasoning capability, the reasoning can over-ride the gene-specified rewards to produce behaviour that is in the interests of the individual, the phenotype, and not the genes, and such behaviour is therefore even much less influenced (not 'determined') by genes.

Emotion and Decision-Making Explained further goes beyond the brain mechanisms of emotion with a treatment (in Chapter 4) of the many different learning processes that become engaged in relation to emotion. The book also includes a formal treatment (in Appendix 1) of reinforcement learning and temporal difference (TD) learning, which are increasingly being used to understand emotion-related learning, as well as its brain mechanisms.

Emotion and Decision-Making Explained goes beyond the brain mechanisms of emotion with a treatment of the functions of affective states in motivated behaviour (including hunger, and sexual behaviour), and indeed proposes a fundamental and simple relation between emotion and motivation. The role of sexual selection in the evolution of affective behaviour is included in Chapter 7.

The book includes findings from the rapidly developing field of neuroeconomics in Chapter 9.

Indeed, in this book, I show how it is now possible to follow processing in the brain from the sensory representation and perception of objects including visual and taste objects that are independent of reward value; to brain regions where reward value (both outcome value and expected value) are represented, which are crucial components of decision-making; to brain mechanisms that actually implement the choice part of the decision-making, with

a mechanism that is common to categorization and decision-making in other brain systems and cortical areas. I believe that this represents a major advance in neuroscience that we are able to understand at the level of mechanisms all of these processes, and to see how they are linked together in the brain to implement much of our behaviour. Moreover, all of this neural understanding is linked to an understanding of the adaptive value of this organization of behaviour, how emotion is a key component, and even how the subjective feeling of pleasure may arise and be related to these processes.

At the same time, *Emotion and Decision-Making Explained* does consider research on how emotion is implemented in the brain, including much new research in the areas of neurophysiology, and functional neuroimaging, neuropsychiatry, and clinical neuropsychology in humans. This treatment of the brain mechanisms of emotion is important not only for providing a basis for understanding disorders of emotion, but also turns out to be important in unravelling the many different ways in which emotions can influence our behaviour, because the different brain mechanisms themselves are being unravelled. The book includes a theory of how the orbitofrontal cortex supports rapid reversals of emotional behaviour, by using a short-term memory network for the current rule which acts in a biased competition mode to influence neurons known to be present in the orbitofrontal cortex. This helps to provide a contrast between the functions of the orbitofrontal cortex and amygdala in emotion. A description of the theory is given in Chapter 4.

Appendix 1 includes a treatment of autoassociation attractor networks that can maintain stable activity in a brain region, and provide a basis for understanding decision-making mechanisms in the brain. Appendix 1 also shows how interacting attractor networks help to provide a foundation for understanding the interactions between mood, and cognition and memory.

The book links to research in psychiatry, with for example discussions of the impulsive behaviour that is a feature of borderline personality disorder, to research in neurology, with for example assessment of the effects on emotion of damage produced by discrete lesions of the human brain, and to research in neuropsychiatry, by introducing recent approaches based on stochastic neurodynamics to the understanding and treatment of schizophrenia and obsessive-compulsive disorder.

Emotion and Decision-Making Explained also goes beyond the brain mechanisms involved in emotion, by addressing (in Chapter 10) emotional feelings, part of the much larger problem of consciousness. One issue developed here is the concept that there is a credit assignment problem if a multiple step plan does not succeed, and that higher-order thoughts provide a solution to this problem. The book also describes many recent functional neuroimaging investigations in which it has been possible to show that the activations of some brain regions are directly correlated with subjective feelings of affective states such as *pleasure*.

Our understanding of emotion, decision-making, and the mechanisms of brain function, described in this book have wider implications, to for example aesthetics, ethics, and the philosophy of mind, and these wider implications are developed in *Neuroculture: On the Implications of Brain Science* (Rolls 2012d). My book *Memory, Attention, and Decision-Making* (Rolls 2008b) shows how some of the neural mechanisms described in this book, and a number of others, provide a unifying computational neuroscience approach to understanding many aspects of brain function, including short-term memory, long-term memory, top-down attention, visual object recognition, and information representation in the brain, as well as decision-making. *Memory, Attention, and Decision-Making* (Rolls 2008b) includes Appendices that may be useful for those wishing an introduction to the computational neuroscience mechanisms involved in many aspects of brain function. *The Noisy Brain: Stochastic Dynamics as a Principle of Brain Function* (Rolls & Deco 2010) describes in detail stochastic dynamics in the brain, how it can be understood with the techniques of theoretical physics, how it contributes

to many aspects of brain function and behaviour, and how it provides new approaches to the cognitive changes that occur with aging, and to psychiatric disorders such as schizophrenia and obsessive-compulsive disorder. The present book replaces *Emotion Explained* (Rolls 2005b) except for *Emotion Explained* Chapter 6 on Thirst, and Chapter 7 on Brain-Stimulation Reward, and both of these Chapters are available at `http://www.oxcns.org`.

It is hoped that this book will be of interest to all those interested in what emotions are, why we have them, how we have them, their disorders, and how we take decisions based on emotions, as well as on rational thinking, and even how we choose between these types of decision-making.

The material in this text is the copyright of Edmund T. Rolls. Part of the material described in the book reflects research performed over many years in collaboration with many colleagues, whose tremendous contributions are warmly appreciated. The contributions of many will be evident from the references cited in the text. In addition, I have benefited enormously from the discussions I have had with a large number of colleagues and friends, many of whom I hope will see areas of the text that they have been able to illuminate. Much of the work described would not have been possible without financial support from a number of sources, particularly the Medical Research Council of the UK, the Human Frontier Science Program, the Wellcome Trust, the McDonnell-Pew Foundation, and the Commission of the European Communities.

The book was typeset by the author in LaTeX using the WinEdt editor.

The cover shows the painting 'Adam and Eve' painted in c. 1528 by Lucas Cranach the Elder (Uffizi Gallery, Florence), which provides an early interpretation of early human emotions, and emotion-related decision-making. This book provides a more recent, scientific, approach to emotions, and to decision-making.

Updates to the publications cited in this book and .pdf files of many papers are available at `http://www.oxcns.org`.

Edmund T. Rolls dedicates this work to the overlapping group: his family, friends, and colleagues: *in salutem praesentium, in memoriam absentium.*

Contents

1 Introduction: the issues **1**

1.1 Introduction 1

1.2 Rewards and punishers 2

1.3 The approaches taken to emotion and motivation 4

1.4 The plan of the book 10

2 The nature of emotion **13**

2.1 Introduction 13

2.2 A theory of emotion 14

2.3 Different emotions 17

2.4 Refinements of the theory of emotion 24

2.5 The classification of emotion 28

2.6 Other theories of emotion 29

2.6.1 The James–Lange and other bodily theories 30

2.6.2 Appraisal theory 33

2.6.3 Dimensional and categorical theories of emotion 35

2.6.4 Other approaches to emotion 35

2.7 Individual differences in emotion, personality, and emotional intelligence 36

2.8 Cognition and emotion 39

2.9 Emotion, motivation, reward, and mood 40

2.10 The concept of emotion 41

2.11 Advantages of the approach to emotion described here (Rolls' theory of emotion) 42

3 The functions of emotion: reward, punishment, and emotion in brain design **45**

3.1 Introduction 45

3.2 Brain design and the functions of emotion 47

3.2.1 Taxes, rewards, and punishers: gene-specified goals for actions, and the flexibility of actions 47

3.2.2 Explicit systems, language, and reinforcement 51

3.2.3 Special-purpose design by an external agent vs evolution by natural selection 51

3.3 Selection of behaviour: cost–benefit 'analysis' of net value 53

3.4 Further functions of emotion 55

3.4.1 Autonomic and endocrine responses 55

3.4.2 Flexibility of behavioural responses 56

3.4.3 Emotional states are motivating 57

3.4.4 Communication 58

3.4.5 Social attachment 60

3.4.6 Separate functions for each different primary reinforcer 61
3.4.7 The mood state can influence the cognitive evaluation of moods or memories 62
3.4.8 Facilitation of memory storage 62
3.4.9 Emotional and mood states are persistent, and help to produce persistent motivation 62
3.4.10 Emotions may trigger memory recall and influence cognitive processing 62
3.5 The functions of emotion in an evolutionary, Darwinian, context 63
3.6 The functions of motivation in an evolutionary, Darwinian, context 65
3.7 Are all goals for action gene-specified? 65

4 The brain mechanisms underlying emotion 67
4.1 Introduction 67
4.2 Overview 67
4.3 Representations of primary reinforcers, i.e. of unlearned value 71
4.3.1 Taste 71
4.3.2 Smell 71
4.3.3 Pleasant and painful touch 72
4.3.4 Visual stimuli 74
4.4 Representing potential secondary reinforcers 76
4.4.1 The requirements of the representation 76
4.4.2 Objects, and not their reward and punishment associations or value, are represented in the inferior temporal visual cortex 80
4.4.3 Object representations 82
4.4.4 Invariant representations of faces and objects in the inferior temporal visual cortex 83
4.4.5 Face expression, gesture and view represented in a population of neurons in the cortex in the superior temporal sulcus 94
4.4.6 The brain mechanisms that build the appropriate view-invariant representations of objects required for learning emotional responses to objects, including faces 94
4.5 The orbitofrontal cortex 95
4.5.1 Historical background 95
4.5.2 Topology 97
4.5.3 Connections 99
4.5.4 Effects of damage to the orbitofrontal cortex 100
4.5.5 Neurophysiology and functional neuroimaging of the orbitofrontal cortex 102
4.5.6 The human orbitofrontal cortex 142
4.5.7 A neurophysiological and computational basis for stimulus–reinforcer association learning and reversal in the orbitofrontal cortex 151
4.5.8 Executive functions of the orbitofrontal cortex 158
4.6 The amygdala 159
4.6.1 Associative processes involved in emotion-related learning 159
4.6.2 Connections of the amygdala 165
4.6.3 Effects of amygdala lesions 167
4.6.4 Neuronal activity in the primate amygdala to reinforcing stimuli 174

4.6.5 Responses of these amygdala neurons to novel stimuli that are reinforcing 181
4.6.6 Neuronal responses in the amygdala to faces 182
4.6.7 Evidence from humans 184
4.6.8 Amygdala summary 188
4.7 The cingulate cortex 189
4.7.1 Introduction and overview of the anterior cingulate cortex 189
4.7.2 Anterior cingulate cortex anatomy and connections 191
4.7.3 Anterior cingulate cortex functional neuroimaging and neuronal activity 191
4.7.4 Anterior cingulate cortex lesion effects 195
4.7.5 Mid-cingulate cortex, the cingulate motor area, and action–outcome learning 196
4.8 Value-related decision-making and medial prefrontal cortex area 10 198
4.8.1 Decision-making between the value of odours 198
4.8.2 Decision-making between the value of thermal somatosensory stimuli 198
4.8.3 Value-related decision-making and the medial prefrontal cortex area 10: further evidence 199
4.9 Insula 200
4.10 Human brain imaging investigations of mood and depression 203
4.11 Output pathways for emotional responses 208
4.11.1 The autonomic and endocrine systems 208
4.11.2 Motor systems for implicit responses, including the basal ganglia 209
4.11.3 Output systems for explicit responses to emotional stimuli 209
4.11.4 Basal forebrain and hypothalamus 209
4.11.5 Basal forebrain cholinergic neurons 210
4.11.6 Noradrenergic neurons 212
4.12 Effects of emotion on cognitive processing and memory 213
4.13 Laterality effects in human emotional processing 218
4.14 Summary 221

5 Food reward value, pleasure, hunger, and appetite 224
5.1 Introduction 224
5.2 Peripheral signals for hunger and satiety 224
5.3 The control signals for hunger and satiety 227
5.3.1 Sensory-specific satiety 227
5.3.2 Gastric distension 232
5.3.3 Duodenal chemosensors 233
5.3.4 Glucostatic hypothesis 233
5.3.5 Hormonal signals related to hunger and satiety, and their effects on the hypothalamus 233
5.3.6 Conditioned appetite and satiety 235
5.4 The brain control of eating and reward 236
5.4.1 The hypothalamus 236
5.4.2 Brain mechanisms for taste reward value 244
5.4.3 Convergence between taste and olfactory processing to represent flavour 256
5.4.4 Brain mechanisms for the reward produced by the odour of food 257

5.4.5 The responses of orbitofrontal cortex taste and olfactory neurons to the sight of food: expected value neurons 262
5.4.6 Functions of the amygdala and temporal cortex in feeding 262
5.4.7 Functions of the orbitofrontal cortex in feeding 266
5.4.8 Output pathways for feeding 269
5.5 Obesity, bulimia, and anorexia 269
5.5.1 Genetic factors 269
5.5.2 Brain processing of the sensory properties and pleasantness of food 270
5.5.3 Food palatability 271
5.5.4 Sensory-specific satiety 272
5.5.5 Fixed meal times, and the availability of food 272
5.5.6 Food saliency, and portion size 273
5.5.7 Energy density of food 273
5.5.8 Eating rate 273
5.5.9 Stress 273
5.5.10 Food craving 273
5.5.11 Energy output 274
5.5.12 Cognitive factors, and attention 274
5.5.13 Compliance with information about risk factors for obesity 274
5.6 Conclusions on reward, affective responses to food, and the control of appetite 275

6 Pharmacology of emotion, reward, and addiction; the basal ganglia 277
6.1 Introduction 277
6.2 Dopamine and reward 279
6.2.1 Dopamine and brain-stimulation reward 279
6.2.2 Self-administration of dopaminergic substances, and addiction 279
6.2.3 Behaviours associated with the release of dopamine 281
6.2.4 The activity of dopaminergic neurons and reward 282
6.3 The basal ganglia 287
6.3.1 Systems-level architecture of the basal ganglia 288
6.3.2 Effects of basal ganglia damage 289
6.3.3 Neuronal activity in the striatum 291
6.3.4 What computations are performed by the basal ganglia? 304
6.3.5 How do the basal ganglia perform their computations? 306
6.3.6 Synthesis on the role of dopamine in reward and addiction 314
6.3.7 Synthesis: emotion, dopamine, reward, punishment, and action selection in the basal ganglia 315
6.4 Opiate reward systems, analgesia, and food reward 317
6.5 Pharmacology of depression in relation to brain systems involved in emotion 318
6.6 Pharmacology of anxiety in relation to brain systems involved in emotion 319
6.7 Cannabinoids 320
6.8 Overview of behavioural selection and output systems involved in emotion 320

7 Sexual behaviour, reward, and brain function; sexual selection of behaviour **323**

7.1 Introduction 323

7.2 The ultimate explanation for the reward value of sex 325

7.3 Mate selection, attractiveness, and love 329
7.3.1 Female preferences 329
7.3.2 Male preferences 331
7.3.3 Pair-bonding, and love 334

7.4 Parental attachment, care, and parent–offspring conflict 335

7.5 Sperm competition and its consequences for sexual behaviour 336

7.6 Concealed ovulation and its consequences for sexual behaviour 343

7.7 Sexual selection of sexual and non-sexual behaviour 344
7.7.1 Sexual selection and natural selection 344
7.7.2 Non-sexual characteristics may be sexually selected for courtship 347

7.8 Individual differences in sexual rewards 348
7.8.1 Overview 349
7.8.2 How might different types of behaviour be produced by natural selection altering the relative reward value of different stimuli in different individuals? 351
7.8.3 How being tuned to different types of reward could help to produce individual differences in sexual behaviour 353

7.9 The neural reward mechanisms that might mediate some aspects of sexual behaviour 355

7.10 Neural basis of sexual behaviour 362

7.11 Conclusion 367

8 Decision-making mechanisms **368**

8.1 Introduction 368

8.2 Decision-making in an attractor network 369
8.2.1 An attractor decision-making network 369
8.2.2 An integrate-and-fire implementation of the attractor network for probabilistic decision-making 371

8.3 Mean-field analysis of the attractor decision-making network 373

8.4 Stability, energy landscapes, and noise 375

8.5 Neurophysiology of vibrotactile decision-making 377

8.6 Probabilistic decision-making by the integrate-and-fire attractor model 380
8.6.1 Integrate-and-fire simulations of decision-making 380
8.6.2 Decision-making on single trials 380
8.6.3 The probabilistic nature of the decision-making 382
8.6.4 Probabilistic decision-making and Weber's law 383
8.6.5 Decision times 386
8.6.6 Finite-size noise effects 387

8.7 Confidence in decisions 390
8.7.1 The model of decision-making 391
8.7.2 Neuronal responses on difficult vs easy trials, and decision confidence 393

8.7.3 Decision times of the neuronal responses 397
8.7.4 Percentage correct 397
8.7.5 Simulation of fMRI signals: haemodynamic convolution of synaptic activity 397
8.7.6 Prediction of the BOLD signals on difficult vs easy decision-making trials 399
8.7.7 Neuroimaging investigations of task difficulty, and confidence 402
8.7.8 Correct decisions vs errors, and confidence 405
8.8 Decisions based on confidence in one's decisions: self-monitoring 417
8.8.1 Decisions about confidence estimates 417
8.8.2 A theory for decisions about confidence estimates 417
8.8.3 Decisions about confidence estimates: neurophysiological evidence 424
8.8.4 Decisions about decisions: self-monitoring 426
8.8.5 Synthesis: decision confidence, noise, neuronal activity, the BOLD signal, and self-monitoring 427
8.9 Perceptual decisions 432
8.10 Comparison with other models of decision-making 433
8.11 Applications and implications of this approach to decision-making 435
8.11.1 Multiple decision-making systems in the brain 435
8.11.2 Distributed decision-making 437
8.11.3 Predicting a decision before the evidence is provided 437
8.11.4 The matching law 439
8.11.5 Symmetry-breaking 439
8.11.6 The evolutionary utility of probabilistic choice 440
8.11.7 Unpredictable behaviour 441
8.11.8 Memory recall 441
8.11.9 Creative thought 442
8.11.10 Decision-making with sequential inputs and with postponed responses 442
8.11.11 Decision-making between the emotional and rational systems 442
8.11.12 Dynamical neuropsychiatry: schizophrenia 442
8.11.13 Dynamical neuropsychiatry: obsessive-compulsive disorder 449
8.11.14 Decision-making, oscillations, and communication through coherence 452

9 Neuroeconomics and decision-making 454
9.1 Introduction 454
9.2 Classical economics 455
9.3 Neoclassical economics 456
9.3.1 Utility functions, WARP, and GARP 456
9.3.2 Expected Utility Theory 457
9.3.3 Random Utility Models 457
9.4 Behavioural economics 458
9.4.1 The Allais paradox 458
9.4.2 Risk seeking over losses 458
9.4.3 Prospect Theory 459
9.5 Neuroeconomics 463
9.5.1 Overview of neuroeconomics 463
9.5.2 A common scale of value for different goods in the orbitofrontal cortex, but no conversion to a common currency 469

9.5.3 Absolute value and relative value are both represented in the orbitofrontal cortex 473
9.5.4 The representation of expected reward value 477
9.5.5 Delay of reward, emotional choice, and rational choice 477
9.5.6 The representation of negative reward prediction error 479
9.5.7 The representation of positive reward prediction error 479
9.5.8 Reward prediction error, temporal difference error, and choice 480
9.5.9 Conclusions 481

10 Emotional feelings and consciousness: a theory of consciousness 483
10.1 Introduction 483
10.2 A Higher-Order Syntactic Thought (HOST) theory of consciousness 484
10.2.1 Multiple routes to action 484
10.2.2 A computational hypothesis of consciousness 487
10.2.3 Adaptive value of processing in the system that is related to consciousness 489
10.2.4 Symbol grounding 490
10.2.5 Qualia 491
10.2.6 Pathways 492
10.2.7 Consciousness and causality 493
10.2.8 Consciousness, a computational system for higher-order syntactic manipulation of symbols, and a commentary or reporting functionality 495
10.3 Selection between conscious vs unconscious decision-making, and free will 496
10.3.1 Dual major routes to action: implicit and explicit 496
10.3.2 The Selfish Gene vs The Selfish Phenotype 502
10.3.3 Decision-making between the implicit and explicit systems 504
10.4 Determinism 504
10.5 Free will 506
10.6 Content and meaning in representations 507
10.7 The causal role of consciousness: a theory of the relation between the mind and the brain 509
10.8 Comparison with other theories of consciousness 511
10.8.1 Higher-order thought theories 511
10.8.2 Oscillations and temporal binding 513
10.8.3 A high neural threshold for information to reach consciousness 514
10.8.4 James–Lange theory and Damasio's somatic marker hypothesis about feelings 515
10.8.5 LeDoux's approach to emotion and consciousness 515
10.8.6 Panksepp's approach to emotion and consciousness 515
10.8.7 Global workspace theories of consciousness 516
10.8.8 Monitoring and consciousness 516

11 Conclusions, and broader issues 518
11.1 Conclusions 518
11.2 Decision-making 525
11.2.1 Selection of mainly autonomic responses, and their classical conditioning 525

11.2.2 Selection of approach or withdrawal, and their classical conditioning; fixed action patterns 526
11.2.3 Selection of fixed stimulus–response habits 526
11.2.4 Selection of arbitrary behaviours to obtain goals, action–outcome learning, and emotional learning 526
11.2.5 The roles of the prefrontal cortex in the selection of action, in decision-making, and in attention 527
11.2.6 Selection of actions by explicit rational thought 533
11.3 Emotion and ethics 535
11.4 Emotion and aesthetics 539
11.5 Close 542

A Neural networks and emotion-related learning 544
A.1 Neurons in the brain, the representation of information, and neuronal learning mechanisms 544
A.1.1 Introduction 544
A.1.2 Neurons in the brain, and their representation in neuronal networks 544
A.1.3 A formalism for approaching the operation of single neurons in a network 545
A.1.4 Synaptic modification 547
A.1.5 Long-Term Potentiation and Long-Term Depression 549
A.1.6 Distributed representations 553
A.2 Pattern association memory 555
A.2.1 Architecture and operation 556
A.2.2 A simple model 558
A.2.3 The vector interpretation 561
A.2.4 Properties 562
A.2.5 Prototype extraction, extraction of central tendency, and noise reduction 565
A.2.6 Speed 565
A.2.7 Local learning rule 566
A.2.8 Implications of different types of coding for storage in pattern associators 571
A.3 Autoassociation memory: attractor networks 572
A.3.1 Architecture and operation 573
A.3.2 Introduction to the analysis of the operation of autoassociation networks 575
A.3.3 Properties 575
A.4 Coupled attractor networks 580
A.5 Reinforcement learning 582
A.5.1 Associative reward–penalty algorithm of Barto and Sutton 583
A.5.2 Error correction or delta rule learning, and classical conditioning 584
A.5.3 Temporal Difference (TD) learning 585

B Decision-making models 589
B.1 Overview of different models of decision-making 589
B.1.1 Sequential-sampling models: sequential probability ratio test, drift-diffusion, and race models 589
B.1.2 Biologically motivated rate models 595

B.1.3 Attractor models 597
B.1.4 Distinguishing model approaches 608
B.2 Synaptic facilitation and sequential decision-making 610
B.3 Synaptic facilitation, graded firing rates, and postponed decisions 612
B.4 The integrate-and-fire formulation used in the model of decision-making 614
B.5 The mean-field approach used in the model of decision-making 616
B.6 The model parameters used in the simulations of decision-making 618

C Glossary 619

References 621

Index 679

D Colour Plates 687