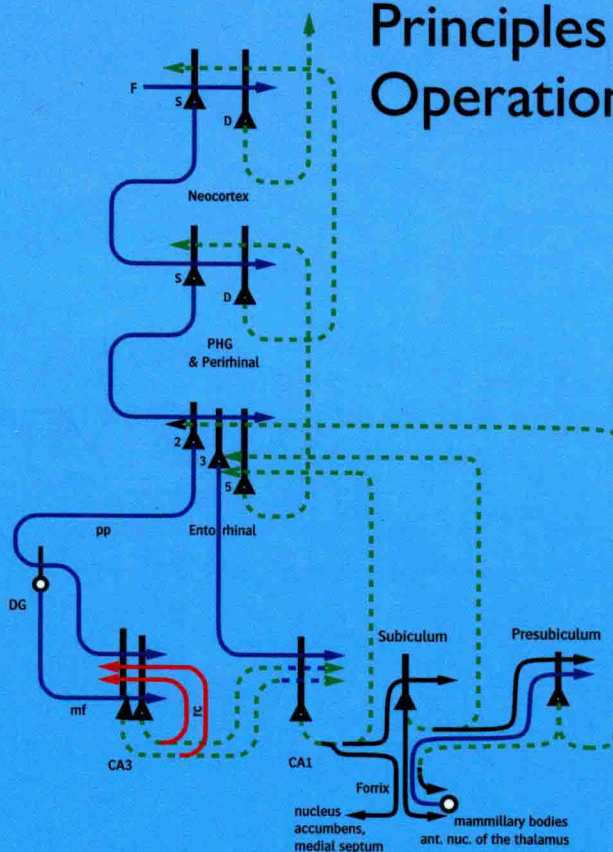


OXFORD

Edmund T. Rolls

# Cerebral Cortex

Principles of Operation



# **Cerebral Cortex**

## Principles of Operation

Edmund T. Rolls

Oxford Centre for Computational Neuroscience, Oxford, UK

**OXFORD**  
UNIVERSITY PRESS



# OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,  
United Kingdom

Oxford University Press is a department of the University of Oxford.  
It furthers the University's objective of excellence in research, scholarship,  
and education by publishing worldwide. Oxford is a registered trade mark of  
Oxford University Press in the UK and in certain other countries

© Edmund Rolls 2016

The moral rights of the author have been asserted

First Edition published in 2016

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in  
a retrieval system, or transmitted, in any form or by any means, without the  
prior permission in writing of Oxford University Press, or as expressly permitted  
by law, by licence or under terms agreed with the appropriate reprographics  
rights organization. Enquiries concerning reproduction outside the scope of the  
above should be sent to the Rights Department, Oxford University Press, at the  
address above

You must not circulate this work in any other form  
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press  
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data

Data available

Library of Congress Control Number: 2016944945

ISBN 978-0-19-878485-2

Printed and bound by

CPI Group (UK) Ltd, Croydon, CR0 4YY

Oxford University Press makes no representation, express or implied, that the  
drug dosages in this book are correct. Readers must therefore always check  
the product information and clinical procedures with the most up-to-date  
published product information and data sheets provided by the manufacturers  
and the most recent codes of conduct and safety regulations. The authors and  
the publishers do not accept responsibility or legal liability for any errors in the  
text or for the misuse or misapplication of material in this work. Except where  
otherwise stated, drug dosages and recommendations are for the non-pregnant  
adult who is not breast-feeding

Links to third party websites are provided by Oxford in good faith and  
for information only. Oxford disclaims any responsibility for the materials  
contained in any third party website referenced in this work.

# Preface

---

The overall aim of this book is to provide insight into the principles of operation of the cerebral cortex. These are key to understanding how we, as humans, function.

There have been few previous attempts to set out some of the important principles of operation of the cortex, and this book is pioneering. I have asked some of the leading investigators in neuroscience about their views on this, and most have not had many well formulated answers or hypotheses. As clear hypotheses are needed in this most important area of 21st century science, how our brains work, I have formulated a set of hypotheses to guide thinking and future research. I present evidence for many of the hypotheses, but at the same time we must all recognise that hypotheses and theory in science are there to be tested, and hopefully refined rather than rejected. Nevertheless, such theories and hypotheses are essential to progress, and it is in this frame of reference that I present the theories, hypotheses, and ideas that I have produced and collected together.

This book focusses on the principles of operation of the cerebral cortex, because at this time it is possible to propose and describe many principles, and many are likely to stand the test of time, and provide a foundation I believe for further developments, even if some need to be changed. In this context, I have not attempted to produce an overall theory of operation of the cerebral cortex, because at this stage of our understanding, such a theory would be incorrect or incomplete. I believe though that many of the principles will be important, and that many will provide the foundations for more complete theories of the operation of the cerebral cortex.

Given that many different principles of operation of the cortex are proposed in this book, with often several principles in each Chapter, the reader may find it convenient to take one Chapter at a time, and think about the issues raised in each Chapter, as the overall enterprise is large. The Highlights sections provided at the end of each Chapter may be useful in helping the reader to appreciate the different principles being considered in each Chapter.

To understand how the cortex works, including how it functions in perception, memory, attention, decision-making, and cognitive functions, 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 different types of memory, perception, attention, and decision-making are implemented by processing in the cerebral cortex.



I emphasize that to understand how memory, perception, attention, decision-making, cognitive functions, and actions are produced in the cortex, we are dealing with large-scale computational systems with interactions between the parts, and that this understanding requires analysis at the computational and global level of the operation of many neurons to perform together a useful function. Understanding at the molecular level is important for helping to understand how these large-scale computational processes are implemented in the brain, but will not by itself give any account of what computations are performed to implement these cognitive functions. Instead, understanding cognitive functions such as object recognition, memory recall, attention, and decision-making requires single neuron data to be closely linked to computational models of how the interactions between large numbers of neurons and many networks of neurons allow these cognitive problems to be solved. The single neuron level is important in this approach, for the single neurons can be thought of as the computational units of the system, and is the level at which the information is exchanged by the spiking activity between the computational elements of the brain. The single neuron level is therefore, because it is the level at which information is communicated between the computing elements of the brain, the fundamental level of information processing, and the level at which the information can be read out (by recording the spiking activity) in order to understand what information is being represented and processed in each brain area.

With its focus on how the brain and especially how the cortex works at the computational neuroscience level, this book is distinct from the many excellent books on neuroscience that describe much evidence about brain structure and function, but do not aim to provide an understanding of how the brain works at the computational level. This book aims to forge an understanding of how some key brain systems may operate at the computational level, so that we can understand how the cortex actually performs some of its complex and necessarily computational functions in memory, perception, attention, decision-making, cognitive functions, and actions.

A test of whether one's understanding is correct is to simulate the processing on a computer, and to show whether the simulation can perform the tasks of cortical systems, and whether the simulation has similar properties to the real cortex. The approach of neural computation leads to a precise definition of how the computation is performed, and to precise and quantitative tests of the theories produced. How memory systems in the cortex work is a paradigm example of this approach, because memory-like operations which involve altered functionality as a result of synaptic modification are at the heart of how many computations in the cortex are performed. It happens that attention and decision-making can be understood in terms of interactions between and fundamental operations in memory systems in the cortex, and therefore it is natural to treat these areas of cognitive neuroscience in this book. The same fundamental concepts based on the operation of neuronal circuitry can be applied to all these functions, as is shown in this book.

One of the distinctive properties of this book is that it links the neural computation approach not only firmly to neuronal neurophysiology, which provides much of the primary data about how the cortex operates, but also to psychophysical studies (for example of attention); to neuropsychological studies of patients with brain damage; and to functional magnetic resonance imaging (fMRI) (and other neuroimaging) approaches. The empirical evidence that is brought to bear is largely from non-human primates and from humans, because of the considerable similarity of their cortical systems.

*In this book, I have not attempted to produce a single computational theory of how the cortex operates. Instead, I have highlighted many different principles of cortical function, most of which are likely to be building blocks of how our cortex operates. The reason for this approach is that many of the principles may well be correct, and useful in understanding how the cortex operates, but some might turn out not to be useful or correct. The aim of this*



*book is therefore to propose some of the fundamental principles of operation of the cerebral cortex, many or most of which will provide a foundation for understanding the operation of the cortex, rather than to produce a single theory of operation of the cortex, which might be disproved if any one of its elements was found to be weak.*

The overall aims of the book are developed further, and the plan of the book is described, in Chapter 1, Section 1.1. Some of the main Principles of Operation of the Cerebral Cortex that I describe can be found in the titles of Chapters 2–22; but in practice, most Chapters include several Principles of Operation, which will appear in the Highlights to each Chapter. Section 26.5 may be useful in addition to the Highlights, for Section 26.5 draws together in a synthesis some of the Principles of Operation of the Cerebral Cortex that are described in the book. Further evidence on how these principles are relevant to the operation of different cortical areas and systems and operate together is provided in Chapters 24–25. In these Chapters, the operation of two major cortical systems, those involved in memory and in visual object recognition, are considered to illustrate how the principles are combined to implement two different key cortical functions. The Appendices provide some of the more formal and quantitative properties of the operation of neuronal systems, and are provided because they provide a route to a deeper understanding on the principles, and to enable the presentation in earlier Chapters to be at a readily approachable level. The Appendices describe many of the building blocks of the neurocomputational approach, and are designed to be useful for teaching. Appendix D describes Matlab software that has been made available with this book to provide simple demonstrations of the operation of some key neuronal networks related to cortical function. The programs are available at <http://www.oxcns.org>.

Part of the material described in the book reflects work performed 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. Especial appreciation is due to Gustavo Deco, Simon M. Stringer, and Alessandro Treves who have contributed greatly in an always interesting and fruitful research collaboration on computational aspects of brain function, and to many neurophysiology and functional neuroimaging colleagues who have contributed to the empirical discoveries that provide the foundation to which the computational neuroscience must always be closely linked, and whose names are cited throughout the text. 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, and the James S. McDonnell Foundation. I am also grateful to many colleagues who I have consulted while writing this book, including Joel Price (Washington University School of Medicine), and Donald Wilson (New York University). Dr Patrick Mills is warmly thanked for his comments on the text. Section 24.3.12 on *ars memoriae* is warmly dedicated to my colleagues at Corpus Christi College, Oxford. The book was typeset by the author using L<sup>A</sup>T<sub>E</sub>X and WinEdt.

The cover includes part of the picture *Pandora* painted in 1896 by J. W. Waterhouse. The metaphor is to look inside the system of the mind and the brain, in order to understand how the brain functions, and thereby better to understand and treat its disorders. The cover also includes an image of the dendritic morphology of excitatory neurons in S1 whisker barrel cortex (Fig. 1.14) (adapted from Marcel Oberlaender, Christiaan P.J. de Kock, Randy M. Bruno, Alejandro Ramirez, Hanno S. Meyer, Vincent J. Dercksen, Moritz Helmstaedter and Bert Sakmann, Cell type-specific three-dimensional structure of thalamocortical circuits in a column of rat vibrissa cortex, *Cerebral Cortex*, 2012, Vol. 22, issue 10, pp. 2375–2391, by permission of Oxford University Press). The cover also includes a diagram of the computational circuitry of the hippocampus by the author (Fig. 24.1). The aim of these second two images is to highlight the importance of moving from the anatomy of the cortex using all the approaches available including neuronal network models that address and

incorporate neurophysiological discoveries to lead to an understanding of how the cortex operates computationally.

Updates to and .pdfs of many of the publications cited in this book are available at <http://www.oxcns.org>. Updates and corrections to the text and notes are also available at <http://www.oxcns.org>.

I dedicate this work to the overlapping group: my family, friends, and colleagues – in salutem praesentium, in memoriam absentium.

# Contents

---

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Principles of operation of the cerebral cortex: introduction and plan	1
1.2	Neurons	4
1.3	Neurons in a network	6
1.4	Synaptic modification	8
1.5	Long-term potentiation and long-term depression	9
1.6	Distributed representations	14
1.6.1	Definitions	14
1.6.2	Advantages of different types of coding	15
1.7	Neuronal network approaches versus connectionism	16
1.8	Introduction to three neuronal network architectures	17
1.9	Systems-level analysis of brain function	18
1.9.1	Ventral cortical visual stream	19
1.9.2	Dorsal cortical visual stream	21
1.9.3	Hippocampal memory system	23
1.9.4	Frontal lobe systems	23
1.9.5	Brodmann areas	24
1.10	The fine structure of the cerebral neocortex	27
1.10.1	The fine structure and connectivity of the neocortex	27
1.10.2	Excitatory cells and connections	27
1.10.3	Inhibitory cells and connections	29
1.10.4	Quantitative aspects of cortical architecture	32
1.10.5	Functional pathways through the cortical layers	34
1.10.6	The scale of lateral excitatory and inhibitory effects, and modules	38
1.11	Highlights	39
<b>2</b>	<b>Hierarchical organization</b>	<b>40</b>
2.1	Introduction	40
2.2	Hierarchical organization in sensory systems	41
2.2.1	Hierarchical organization in the ventral visual system	41
2.2.2	Hierarchical organization in the dorsal visual system	46
2.2.3	Hierarchical organization of taste processing	48
2.2.4	Hierarchical organization of olfactory processing	57
2.2.5	Hierarchical multimodal convergence of taste, olfaction, and vision	59
2.2.6	Hierarchical organization of auditory processing	64
2.3	Hierarchical organization of reward value processing	67
2.4	Hierarchical organization of connections to the frontal lobe for short-term memory	68
2.5	Highlights	69
<b>3</b>	<b>Localization of function</b>	<b>72</b>
3.1	Hierarchical processing	72



3.2	Short-range neocortical recurrent collaterals	72
3.3	Topographic maps	72
3.4	Modularity	72
3.5	Lateralization of function	73
3.6	Ventral and dorsal cortical areas	73
3.7	Highlights	74
<b>4</b>	<b>Recurrent collateral connections and attractor networks</b>	<b>75</b>
4.1	Introduction	75
4.2	Attractor networks implemented by the recurrent collaterals	75
4.3	Evidence for attractor networks implemented by recurrent collateral connections	76
4.3.1	Short-term Memory	77
4.3.2	Long-term Memory	80
4.3.3	Decision-Making	80
4.4	The storage capacity of attractor networks	80
4.5	A global attractor network in hippocampal CA3, but local in neocortex	81
4.6	The speed of operation of cortical attractor networks	83
4.7	Dilution of recurrent collateral cortical connectivity	83
4.8	Self-organizing topographic maps in the neocortex	85
4.9	Attractors formed by forward and backward connections between cortical areas?	85
4.10	Interacting attractor networks	86
4.11	Highlights	90
<b>5</b>	<b>The noisy cortex: stochastic dynamics, decisions, and memory</b>	<b>91</b>
5.1	Reasons why the brain is inherently noisy and stochastic	91
5.2	Attractor networks, energy landscapes, and stochastic neurodynamics	95
5.3	A multistable system with noise	98
5.4	Stochastic dynamics and the stability of short-term memory	101
5.4.1	Analysis of the stability of short-term memory	103
5.4.2	Stability and noise in a model of short-term memory	104
5.5	Long-term memory recall	106
5.6	Stochastic dynamics and probabilistic decision-making in an attractor network	106
5.6.1	Decision-making in an attractor network	107
5.6.2	Theoretical framework: a probabilistic attractor network	107
5.6.3	Stationary multistability analysis: mean-field	110
5.6.4	Integrate-and-fire simulations of decision-making: spiking dynamics	112
5.6.5	Reaction times of the neuronal responses	116
5.6.6	Percentage correct	117
5.6.7	Finite-size noise effects	117
5.6.8	Comparison with neuronal data during decision-making	119
5.6.9	Testing the model of decision-making with human functional neuroimaging	122
5.6.10	Decisions based on confidence in one's decisions: self-monitoring	129
5.6.11	Decision-making with multiple alternatives	131
5.6.12	The matching law	132
5.6.13	Comparison with other models of decision-making	132
5.7	Perceptual decision-making and rivalry	134
5.8	Symmetry-breaking	135

5.9	The evolutionary utility of probabilistic choice	135
5.10	Selection between conscious vs unconscious decision-making, and free will	136
5.11	Creative thought	137
5.12	Unpredictable behaviour	138
5.13	Predicting a decision before the evidence is applied	138
5.14	Highlights	140
<b>6</b>	<b>Attention, short-term memory, and biased competition</b>	<b>141</b>
6.1	Bottom-up attention	141
6.2	Top-down attention – biased competition	143
6.2.1	The biased competition hypothesis	143
6.2.2	Biased competition – single neuron studies	145
6.2.3	Non-spatial attention	147
6.2.4	Biased competition – fMRI	149
6.2.5	A basic computational module for biased competition	149
6.2.6	Architecture of a model of attention	150
6.2.7	Simulations of basic experimental findings	154
6.2.8	Object recognition and spatial search	158
6.2.9	The neuronal and biophysical mechanisms of attention	163
6.2.10	'Serial' vs 'parallel' attentional processing	167
6.3	Top-down attention – biased activation	171
6.3.1	Selective attention can selectively activate different cortical areas	171
6.3.2	Sources of the top-down modulation of attention	173
6.3.3	Granger causality used to investigate the source of the top-down biasing	174
6.3.4	Top-down cognitive modulation	175
6.3.5	A top-down biased activation model of attention	178
6.4	Conclusions	181
6.5	Highlights	184
<b>7</b>	<b>Diluted connectivity</b>	<b>186</b>
7.1	Introduction	186
7.2	Diluted connectivity and the storage capacity of attractor networks	187
7.2.1	The autoassociative or attractor network architecture being studied	187
7.2.2	The storage capacity of attractor networks with diluted connectivity	188
7.2.3	The network simulated	190
7.2.4	The effects of diluted connectivity on the capacity of attractor networks	192
7.2.5	Synthesis of the effects of diluted connectivity in attractor networks	197
7.3	The effects of dilution on the capacity of pattern association networks	198
7.4	The effects of dilution on the performance of competitive networks	201
7.4.1	Competitive Networks	201
7.4.2	Competitive networks without learning but with diluted connectivity	202
7.4.3	Competitive networks with learning and with diluted connectivity	203
7.4.4	Competitive networks with learning and with full (undiluted) connectivity	205
7.4.5	Overview and implications of diluted connectivity in competitive networks	206
7.5	The effects of dilution on the noise in attractor networks	207
7.6	Highlights	207
<b>8</b>	<b>Coding principles</b>	<b>209</b>
8.1	Types of encoding	209



8.2	Place coding with sparse distributed firing rate representations	210
8.2.1	Reading the code used by single neurons	210
8.2.2	Understanding the code provided by populations of neurons	214
8.3	Synchrony, coherence, and binding	221
8.4	Principles by which the representations are formed	222
8.5	Information encoding in the human cortex	223
8.6	Highlights	226
<b>9</b>	<b>Synaptic modification for learning</b>	<b>227</b>
9.1	Introduction	227
9.2	Associative synaptic modification implemented by long-term potentiation	227
9.3	Forgetting in associative neural networks, and memory reconsolidation	228
9.3.1	Forgetting	228
9.3.2	Factors that influence synaptic modification	230
9.3.3	Reconsolidation	232
9.4	Spike-timing dependent plasticity	233
9.5	Long-term synaptic depression in the cerebellar cortex	233
9.6	Reward prediction error learning	234
9.6.1	Blocking and delta-rule learning	234
9.6.2	Dopamine neuron firing and reward prediction error learning	234
9.7	Highlights	240
<b>10</b>	<b>Synaptic and neuronal adaptation and facilitation</b>	<b>241</b>
10.1	Mechanisms for neuronal adaptation and synaptic depression and facilitation	241
10.1.1	Sodium inactivation leading to neuronal spike-frequency adaptation	241
10.1.2	Calcium activated hyper-polarizing potassium current	242
10.1.3	Short-term synaptic depression and facilitation	243
10.2	Short-term depression of thalamic input to the cortex	244
10.3	Relatively little adaptation in primate cortex when it is operating normally	244
10.4	Acetylcholine, noradrenaline, and other modulators of adaptation and facilitation	247
10.4.1	Acetylcholine	247
10.4.2	Noradrenergic neurons	248
10.5	Synaptic depression and sensory-specific satiety	249
10.6	Neuronal and synaptic adaptation, and the memory for sequential order	250
10.7	Destabilization of short-term memory by adaptation or synaptic depression	250
10.8	Non-reward computation in the orbitofrontal cortex using synaptic depression	251
10.9	Synaptic facilitation and a multiple-item short-term memory	253
10.10	Synaptic facilitation in decision-making	253
10.11	Highlights	254
<b>11</b>	<b>Backprojections in the neocortex</b>	<b>255</b>
11.1	Architecture	255
11.2	Learning	257
11.3	Recall	258
11.4	Semantic priming	259
11.5	Top-down Attention	259
11.6	Autoassociative storage, and constraint satisfaction	261

11.7	Highlights	261
<b>12</b>	<b>Memory and the hippocampus</b>	<b>262</b>
12.1	Introduction	262
12.2	Hippocampal circuitry and connections	262
12.3	The hippocampus and episodic memory	262
12.4	Autoassociation in the CA3 network for episodic memory	263
12.5	The dentate gyrus as a pattern separation mechanism, and neurogenesis	265
12.6	Rodent place cells vs primate spatial view cells	265
12.7	Backprojections, and the recall of information from the hippocampus to neocortex	266
12.8	Subcortical structures connected to the hippocampo-cortical memory system	267
12.9	Highlights	267
<b>13</b>	<b>Limited neurogenesis in the adult cortex</b>	<b>269</b>
13.1	No neurogenesis in the adult neocortex	269
13.2	Limited neurogenesis in the adult hippocampal dentate gyrus	269
13.3	Neurogenesis in the chemosensing receptor systems	270
13.4	Highlights	271
<b>14</b>	<b>Invariance learning and vision</b>	<b>272</b>
14.1	Hierarchical cortical organization with convergence	272
14.2	Feature combinations	272
14.3	Sparse distributed representations	273
14.4	Self-organization by feedforward processing without a teacher	273
14.5	Learning guided by the statistics of the visual inputs	274
14.6	Bottom up saliency	275
14.7	Lateral interactions shape receptive fields	276
14.8	Top-down selective attention vs feedforward processing	277
14.9	Topological maps to simplify connectivity	278
14.10	Biologically decodable output representations	279
14.11	Highlights	279
<b>15</b>	<b>Emotion, motivation, reward value, pleasure, and their mechanisms</b>	<b>281</b>
15.1	Emotion, reward value, and their evolutionary adaptive utility	281
15.2	Motivation and reward value	283
15.3	Principles of cortical design for emotion and motivation	283
15.4	Objects are first represented independently of reward value	284
15.5	Specialized systems for face identity and expression processing in primates	286
15.6	Unimodal processing to the object level before multimodal convergence	287
15.7	A common scale for reward value	287
15.8	Sensory-specific satiety	287
15.9	Economic value is represented in the orbitofrontal cortex	288
15.10	Neuroeconomics vs classical microeconomics	288
15.11	Output systems influenced by orbitofrontal cortex reward value representations	289
15.12	Decision-making about rewards in the anterior orbitofrontal cortex	291



15.13	Probabilistic emotion-related decision-making	292
15.14	Non-reward, error, neurons in the orbitofrontal cortex	292
15.15	Reward reversal learning in the orbitofrontal cortex	296
15.16	Dopamine neurons and emotion	301
15.17	The explicit reasoning system vs the emotional system	301
15.18	Pleasure	302
15.19	Personality relates to differences in sensitivity to rewards and punishers	302
15.20	Highlights	303
<b>16</b>	<b>Noise in the cortex, stability, psychiatric disease, and aging</b>	<b>305</b>
16.1	Stochastic noise, attractor dynamics, and schizophrenia	305
16.1.1	Introduction	305
16.1.2	A dynamical systems hypothesis of the symptoms of schizophrenia	307
16.1.3	The depth of the basins of attraction: mean-field flow analysis	308
16.1.4	Decreased stability produced by reduced NMDA conductances	309
16.1.5	Increased distractibility produced by reduced NMDA conductances	311
16.1.6	Synthesis: network instability and schizophrenia	312
16.2	Stochastic noise, attractor dynamics, and obsessive-compulsive disorder	316
16.2.1	Introduction	316
16.2.2	A hypothesis about obsessive-compulsive disorder	317
16.2.3	Glutamate and increased depth of the basins of attraction	319
16.2.4	Synthesis on obsessive-compulsive disorder	322
16.3	Stochastic noise, attractor dynamics, and depression	325
16.3.1	Introduction	325
16.3.2	A non-reward attractor theory of depression	328
16.3.3	Evidence consistent with the theory	329
16.3.4	Relation to other brain systems implicated in depression	331
16.3.5	Implications for treatments	332
16.3.6	Mania and bipolar disorder	333
16.4	Stochastic noise, attractor dynamics, and aging	335
16.4.1	NMDA receptor hypofunction	335
16.4.2	Dopamine	338
16.4.3	Impaired synaptic modification	338
16.4.4	Cholinergic function and memory	339
16.5	Highlights	343
<b>17</b>	<b>Syntax and Language</b>	<b>345</b>
17.1	Neurodynamical hypotheses about language and syntax	345
17.1.1	Binding by synchrony?	345
17.1.2	Syntax using a place code	346
17.1.3	Temporal trajectories through a state space of attractors	347
17.1.4	Hypotheses about the implementation of language in the cerebral cortex	347
17.2	Tests of the hypotheses – a model	351
17.2.1	Attractor networks with stronger forward than backward connections	351
17.2.2	The operation of a single attractor network module	353
17.2.3	Spike frequency adaptation mechanism	355
17.3	Tests of the hypotheses – findings with the model	355
17.3.1	A production system	355
17.3.2	A decoding system	356
17.4	Evaluation of the hypotheses	359

17.5	Highlights	363
<b>18</b>	<b>Evolutionary trends in cortical design and principles of operation</b>	<b>364</b>
18.1	Introduction	364
18.2	Different types of cerebral neocortex: towards a computational understanding	364
18.2.1	Neocortex or isocortex	365
18.2.2	Olfactory (pyriform) cortex	371
18.2.3	Hippocampal cortex	374
18.3	Addition of areas in the neocortical hierarchy	376
18.4	Evolution of the orbitofrontal cortex	378
18.5	Evolution of the taste and flavour system	379
18.5.1	Principles	379
18.5.2	Taste processing in rodents	380
18.6	Evolution of the temporal lobe cortex	381
18.7	Evolution of the frontal lobe cortex	382
18.8	Highlights	382
<b>19</b>	<b>Genetics and self-organization build the cortex</b>	<b>385</b>
19.1	Introduction	385
19.2	Hypotheses about the genes that build cortical neural networks	386
19.3	Genetic selection of neuronal network parameters	390
19.4	Simulation of the evolution of neural networks using a genetic algorithm	391
19.4.1	The neural networks	391
19.4.2	The specification of the genes	392
19.4.3	The genetic algorithm, and general procedure	397
19.4.4	Pattern association networks	398
19.4.5	Autoassociative networks	400
19.4.6	Competitive networks	400
19.5	Evaluation of the gene-based evolution of single-layer networks	401
19.6	The gene-based evolution of multi-layer cortical systems	403
19.7	Highlights	404
<b>20</b>	<b>Cortex versus basal ganglia design for selection</b>	<b>406</b>
20.1	Systems-level architecture of the basal ganglia	406
20.2	What computations are performed by the basal ganglia?	408
20.3	How do the basal ganglia perform their computations?	410
20.4	Comparison of selection in the basal ganglia and cerebral cortex	413
20.5	Highlights	415
<b>21</b>	<b>Sleep and Dreaming</b>	<b>416</b>
21.1	Is sleep necessary for cortical function?	416
21.2	Is sleep involved in memory consolidation?	417
21.3	Dreams	418
21.4	Highlights	419
<b>22</b>	<b>Which cortical computations underlie consciousness?</b>	<b>420</b>
22.1	Introduction	420



22.2	A Higher-Order Syntactic Thought (HOST) theory of consciousness	421
22.2.1	Multiple routes to action	421
22.2.2	A computational hypothesis of consciousness	423
22.2.3	Adaptive value of processing that is related to consciousness	425
22.2.4	Symbol grounding	426
22.2.5	Qualia	428
22.2.6	Pathways	429
22.2.7	Consciousness and causality	430
22.2.8	Consciousness and higher-order syntactic thoughts	431
22.3	Selection between conscious vs unconscious decision-making systems	432
22.3.1	Dual major routes to action: implicit and explicit	432
22.3.2	The Selfish Gene vs The Selfish Phenotype	439
22.3.3	Decision-making between the implicit and explicit systems	440
22.4	Determinism	441
22.5	Free will	442
22.6	Content and meaning in representations	443
22.7	The causal role of consciousness and the relation between the mind and the brain	445
22.8	Comparison with other theories of consciousness	447
22.8.1	Higher-order thought theories	447
22.8.2	Oscillations and temporal binding	449
22.8.3	A high neural threshold for information to reach consciousness	450
22.8.4	James–Lange theory and Damasio's somatic marker hypothesis	451
22.8.5	LeDoux's approach to emotion and consciousness	451
22.8.6	Panksepp's approach to emotion and consciousness	452
22.8.7	Global workspace theories of consciousness	452
22.8.8	Monitoring and consciousness	452
22.9	Highlights	453
<b>23</b>	<b>Cerebellar cortex</b>	<b>455</b>
23.1	Introduction	455
23.2	Architecture of the cerebellum	456
23.2.1	The connections of the parallel fibres onto the Purkinje cells	456
23.2.2	The climbing fibre input to the Purkinje cell	457
23.2.3	The mossy fibre to granule cell connectivity	457
23.3	Modifiable synapses of parallel fibres onto Purkinje cell dendrites	460
23.4	The cerebellar cortex as a perceptron	460
23.5	Highlights: differences between cerebral and cerebellar cortex microcircuitry	461
<b>24</b>	<b>The hippocampus and memory</b>	<b>463</b>
24.1	Introduction	463
24.2	Systems-level functions of the hippocampus	464
24.2.1	Systems-level anatomy	465
24.2.2	Evidence from the effects of damage to the hippocampus	467
24.2.3	The necessity to recall information from the hippocampus	468
24.2.4	Systems-level neurophysiology of the primate hippocampus	470
24.2.5	Head direction cells in the presubiculum	478
24.2.6	Perirhinal cortex, recognition memory, and long-term familiarity memory	479
24.3	A theory of the operation of hippocampal circuitry as a memory system	486
24.3.1	Hippocampal circuitry	487
24.3.2	Entorhinal cortex	488

24.3.3	CA3 as an autoassociation memory	490
24.3.4	Dentate granule cells	509
24.3.5	CA1 cells	515
24.3.6	Recoding in CA1 to facilitate retrieval to the neocortex	515
24.3.7	Backprojections to the neocortex, memory recall, and consolidation	520
24.3.8	Backprojections to the neocortex – quantitative aspects	523
24.3.9	Simulations of hippocampal operation	526
24.3.10	The learning of spatial view and place cell representations	528
24.3.11	Linking the inferior temporal visual cortex to spatial view and place cells	529
24.3.12	A scientific theory of the art of memory: <i>scientia artis memoriae</i>	531
24.4	Tests of the theory of hippocampal cortex operation	531
24.4.1	Dentate gyrus (DG) subregion of the hippocampus	531
24.4.2	CA3 subregion of the hippocampus	535
24.4.3	CA1 subregion of the hippocampus	542
24.5	Evaluation of the theory of hippocampal cortex operation	546
24.5.1	Tests of the theory by hippocampal system subregion analyses	546
24.5.2	Comparison with other theories of hippocampal function	548
24.6	Highlights	552
<b>25</b>	<b>Invariant visual object recognition learning</b>	<b>554</b>
25.1	Introduction	554
25.2	Invariant representations of faces and objects in the inferior temporal visual cortex	555
25.2.1	Processing to the inferior temporal cortex in the primate visual system	555
25.2.2	Translation invariance and receptive field size	556
25.2.3	Reduced translation invariance in natural scenes	557
25.2.4	Size and spatial frequency invariance	560
25.2.5	Combinations of features in the correct spatial configuration	561
25.2.6	A view-invariant representation	562
25.2.7	Learning in the inferior temporal cortex	565
25.2.8	Distributed encoding	568
25.2.9	Face expression, gesture, and view	572
25.2.10	Specialized regions in the temporal cortical visual areas	572
25.3	Approaches to invariant object recognition	576
25.3.1	Feature spaces	577
25.3.2	Structural descriptions and syntactic pattern recognition	578
25.3.3	Template matching and the alignment approach	580
25.3.4	Invertible networks that can reconstruct their inputs	581
25.3.5	Feature hierarchies	582
25.4	Hypotheses about object recognition mechanisms	582
25.5	Computational issues in feature hierarchies	586
25.5.1	The architecture of VisNet	587
25.5.2	Initial experiments with VisNet	596
25.5.3	The optimal parameters for the temporal trace used in the learning rule	603
25.5.4	Different forms of the trace learning rule, and error correction	604
25.5.5	The issue of feature binding, and a solution	612
25.5.6	Operation in a cluttered environment	624
25.5.7	Learning 3D transforms	631
25.5.8	Capacity of the architecture, and an attractor implementation	636
25.5.9	Vision in natural scenes – effects of background versus attention	643
25.5.10	The representation of multiple objects in a scene	651
25.5.11	Learning invariant representations using spatial continuity	653
25.5.12	Lighting invariance	654

25.5.13	Invariant global motion in the dorsal visual system	656
25.5.14	Deformation-invariant object recognition	656
25.5.15	Learning invariant representations of scenes and places	657
25.5.16	Finding and recognising objects in natural scenes	659
25.6	Further approaches to invariant object recognition	663
25.6.1	Other types of slow learning	663
25.6.2	HMAX	663
25.6.3	Sigma-Pi synapses	668
25.6.4	Deep learning	668
25.7	Visuo-spatial scratchpad memory, and change blindness	669
25.8	Processes involved in object identification	670
25.9	Highlights	671
<b>26</b>	<b>Synthesis</b>	<b>674</b>
26.1	Principles of cortical operation, not a single theory	674
26.2	Levels of explanation, and the mind-brain problem	674
26.3	Brain computation compared to computation on a digital computer	676
26.4	Understanding how the brain works	681
26.5	Synthesis on principles of operation of the cerebral cortex	683
26.5.1	Hierarchical organization	683
26.5.2	Localization of function	684
26.5.3	Recurrent collaterals and attractor networks	684
26.5.4	The noisy cortex	685
26.5.5	Top-down attention	685
26.5.6	Diluted connectivity	685
26.5.7	Sparse distributed graded firing rate encoding	685
26.5.8	Synaptic modification	686
26.5.9	Adaptation and facilitation	686
26.5.10	Backprojections	686
26.5.11	Neurogenesis	687
26.5.12	Binding and syntax	687
26.5.13	Evolution of the cerebral cortex	687
26.5.14	Genetic specification of cortical design	687
26.5.15	The cortical systems for emotion	688
26.5.16	Memory systems	688
26.5.17	Visual cortical processing for invariant visual object recognition	689
26.5.18	Cortical lamination, operation, and evolution	689
26.6	Highlights	692
<b>A</b>	<b>Introduction to linear algebra for neural networks</b>	<b>694</b>
A.1	Vectors	694
A.1.1	The inner or dot product of two vectors	694
A.1.2	The length of a vector	695
A.1.3	Normalizing the length of a vector	696
A.1.4	The angle between two vectors: the normalized dot product	696
A.1.5	The outer product of two vectors	697
A.1.6	Linear and non-linear systems	698
A.1.7	Linear combinations, linear independence, and linear separability	699
A.2	Application to understanding simple neural networks	700
A.2.1	Capability and limitations of single-layer networks	701
A.2.2	Non-linear networks: neurons with non-linear activation functions	703