

New Horizons in the Neuroscience of Consciousness

Edited by

Elaine Perry, Daniel Collerton,
Fiona LeBeau and Heather Ashton

Advances in Consciousness Research 79



John Benjamins Publishing Company

New Horizons in the Neuroscience of Consciousness

Advances in Consciousness Research (AiCR)

Provides a forum for scholars from different scientific disciplines and fields of knowledge who study consciousness in its multifaceted aspects. Thus the Series includes (but is not limited to) the various areas of cognitive science, including cognitive psychology, brain science, philosophy and linguistics. The orientation of the series is toward developing new interdisciplinary and integrative approaches for the investigation, description and theory of consciousness, as well as the practical consequences of this research for the individual in society.

From 1999 the Series consists of two subseries that cover the most important types of contributions to consciousness studies:

Series A: Theory and Method. Contributions to the development of theory and method in the study of consciousness; Series B: Research in Progress.

Experimental, descriptive and clinical research in consciousness.

This book is a contribution to Series B.

Editor

Maxim I. Stamenov

Bulgarian Academy of Sciences

Editorial Board

David J. Chalmers

Australian National University

Axel Cleeremans

Université Libre de Bruxelles

Gordon G. Globus

University of California Irvine

Christof Koch

California Institute of Technology

Stephen M. Kosslyn

Harvard University

Steven Laureys

University of Liège

George Mandler

University of California at San Diego

John R. Searle

University of California at Berkeley

Petra Stoerig

Universität Düsseldorf

Volume 79

New Horizons in the Neuroscience of Consciousness

Edited by Elaine Perry, Daniel Collerton, Fiona LeBeau and Heather Ashton

New Horizons in the Neuroscience of Consciousness

Edited by

Elaine Perry

Daniel Collerton

Fiona LeBeau

Heather Ashton

Newcastle University

John Benjamins Publishing Company

Amsterdam / Philadelphia



The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

New horizons in the neuroscience of consciousness / edited by Elaine Perry ... [et al.].

p. cm. (Advances in Consciousness Research, ISSN 1381-589X ; v. 79)

Includes bibliographical references and index.

1. Consciousness. 2. Neurosciences. I. Perry, E. K. (Elaine K.)

QP411.N49 2010

612.8'2--dc22

2010018686

ISBN 978 90 272 5215 9 (Hb ; alk. paper)

ISBN 978 90 272 8804 2 (Eb)

© 2010 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands

John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

List of contributors	IX
Prologue	XIII
Poem	XXV

Section I. Neuronal mechanisms

The slow cortical potential hypothesis on consciousness <i>Biyu J. He and Marcus E. Raichle</i>	3
Distinct characteristics of conscious experience are met by large-scale neuronal synchronization <i>Lucia Melloni and Wolf Singer</i>	17
Gamma oscillations and the cellular components of consciousness? <i>Fiona E. N. LeBeau</i>	29
Dopamine modulation of decision making processes <i>Anthony A. Grace</i>	39
Undercurrents of consciousness: The endocannabinoid system <i>Heather Ashton</i>	53
Disconnecting consciousness: The neuroscience of general anaesthesia <i>Lucas Campos, Adrian Pichurko and George A. Mashour</i>	65
Consciousness and neural time travel <i>Michael E. Hasselmo</i>	73

Section II. Psychological processes

Consciousness and the relation between implicit and explicit memory <i>Signy Sheldon and Morris Moscovitch</i>	83
Two varieties of unconscious processes <i>Stan Franklin and Bernard J. Baars</i>	91

Operating characteristics and awareness <i>Colin W. G. Clifford, Justin A. Harris and Ehsan Arabzadeh</i>	103
Noise in the brain, decision-making, determinism, free will, and consciousness <i>Edmund T. Rolls</i>	113
Social consciousness <i>Kai Vogeley</i>	121
Consciousness and language: A processing perspective <i>Michael Sharwood Smith and John Truscott</i>	129
Cognitive illusions: From magic to science <i>Gustav Kuhn</i>	139
Dreaming as a model system for consciousness research <i>Antti Revonsuo and Katja Valli</i>	149
Lucid dreaming and the bimodality of consciousness <i>Allan Hobson and Ursula Voss</i>	155
Section III. Psychopathologies and therapies	
Why depression feels bad <i>Mark Solms and Jaak Panksepp</i>	169
Dementia and the boundary between conscious and nonconscious awareness <i>Daniel Collerton</i>	179
Consciousness as the spin-off and schizophrenia as the price of language <i>Timothy J. Crow</i>	187
Consciousness and psychosis associated with schizophrenia: The role of Cornu Ammonis Region 3 <i>Ralf-Peter Behrendt</i>	201
The visual unconscious. Perspectives from the Charles Bonnet Syndrome <i>Dominic H. ffytche</i>	215
Believing is hearing is believing: The reciprocal nature of consciousness <i>William Sedley</i>	227

Dreaming as a physiological psychosis: Connecting states of consciousness	239
<i>Armando D'Agostino, Ivan Limosani and Silvio Scarone</i>	
Conscious awareness versus optimistic beliefs in recreational Ecstasy/MDMA users	249
<i>Andrew Parrott</i>	
Conscious and unconscious placebo responses: How the ritual of the therapeutic act changes the patient's brain	259
<i>Fabrizio Benedetti</i>	
Section IV. Expanding boundaries	
The paradoxes of creativity	271
<i>Ashish Ranpura and Mark F. Lythgoe</i>	
Potential contributions of research on meditation to the neuroscience of consciousness	281
<i>Antoine Lutz</i>	
Self-induced altered states of consciousness	289
<i>Bangalore N. Gangadhar and Naren P. Rao</i>	
Beyond the boundaries of the brain	301
<i>Dean Radin</i>	
Plants of the gods and shamanic journeys	309
<i>Elaine K. Perry and Valerie Laws</i>	
Index	325

List of contributors

Ehsan Arabzadeh

School of Psychology,
University of New South Wales,
Sydney, NSW, Australia.
e-mail: ehsan@unsw.edu.au

Heather Ashton

Department of Psychiatry,
Newcastle University,
Newcastle-upon-Tyne, U.K.
e-mail: c.h.ashton@ncl.ac.uk

Bernard J. Baars

The Neurosciences Institute,
San Diego, California, U.S.A.
e-mail: bbaars@nsi.edu

Ralf-Peter Behrendt

Older People's Services,
The Retreat Hospital, York, U.K.
e-mail: rp.behrendt@btinternet.com

Fabrizio Benedetti

Department of Neuroscience,
University of Turin Medical School,
Turin, Italy.
e-mail: fabrizio.benedetti@unito.it

Lucas Campos

Colin W. G. Clifford

School of Psychology, University of Sydney,
Sydney, NSW, Australia.
e-mail: colin.clifford@sydney.edu.au

Daniel Collerton

Northumberland, Tyne and Wear NHS Trust,
Newcastle University,
Newcastle-upon-Tyne, U.K.
e-mail: daniel.collerton@ncl.ac.uk

Timoth J. Crow

SANE POWIC, Warneford Hospital,
Oxford, U.K.
e-mail: tim.crow@psych.ox.ac.uk

Armando D'Agostino

Department of Medicine,
Surgery and Dentistry, University of Milan,
Milan, Italy.
e-mail: arma.dagostino@gmail.com

Stan Franklin

Department of Computer Science
and Institute for Intelligent Systems,
The University of Memphis,
Memphis, TN, U.S.A.
e-mail: franklin@memphis.edu

Dominic H. ffytche

Section of Old Age Psychiatry
and Centre for Neuroimaging Sciences,
Institute of Psychiatry, London, U.K.
e-mail: d.ffytche@iop.kcl.ac.uk

Bangalore N. Gangadhar

Department of Psychiatry,
National Institute of Mental Health
and Neuroscience,
Bangalore, India.
e-mail: kalyanybg@yahoo.com

Anthony A. Grace

Department of Neuroscience,
University of Pittsburg, PA, U.S.A.
e-mail: graceaa@pitt.edu

Justin A. Harris

School of Psychology,
University of Sydney,
Sydney, NSW, Australia.
e-mail: justin.harris@sydney.edu.au

Michael E. Hasselmo

Department of Psychology
and Program in Neuroscience,
Boston University,
Boston, MA, U.S.A.
e-mail: hasselmo@bu.edu.

Biyu J. He

Mallinckrodt Institute of Radiology,
Washington University School of Medicine,
St. Louis, MO, U.S.A.
e-mail: biyu.jade.he@gmail.com

Allan Hobson

Beth Israel Deaconess Medical Centre
and Harvard Medical Centre,
Boston, MA, U.S.A.
allan_hobson@hms.harvard.edu

Gustav Kuhn

Department of Psychiatry,
Brunel University,
Uxbridge, U.K.
e-mail: www.gustavkuhn.com

Valerie Laws

Writer in Residence,
Institute for Ageing and Health,
Newcastle University,
Newcastle-upon-Tyne, U.K.
e-mail: valaws@yahoo.co.uk

Fiona E. N. LeBeau

Institute of Neuroscience,
Newcastle University,
Newcastle-upon-Tyne, U.K.
e-mail: fiona.lebeau@ncl.ac.uk

Ivan Limosani

Department of Mental Health,
San Paolo Hospital,
Milan, Italy.
e-mail: ilimosani@gmail.com

Antoine Lutz

Laboratory for Functional Brain
Imaging and Behaviour,
Waisman Centre, University
of Wisconsin-Madison, WI, U.S.A.
alutz@facstaff.wisc.edu

Mark F. Lythgoe

Centre for Advanced Biomedical Imaging,
University College London,
London, U.K.
e-mail: mlythgoe@ich.ucl.ac.uk

George A. Mashour

Department of Anaesthesiology,
University of Michigan Medical School,
MI, U.S.A.
e-mail: gmashour@med.umich.edu

Lucia Melloni

Department of Neurophysiology,
Max Plank Institute for Brain Research,
Frankfurt, Germany.
e-mail: lmelloni@gmail.com

Morris Moscovitch

Department of Psychology,
University of Toronto,
Toronto, Canada.
e-mail: momos@psych.utoronto.ca

Jaak Panksepp

Department of Psychology,
University of Cape Town,
South Africa.
e-mail: jpanksepp@vetmed.wsu.edu

Andrew C. Parrott

Department of Psychology,
Swansea University,
Swansea, U.K.
e-mail: a.c.parrott@swansea.ac.uk

Elaine K. Perry

Institute for Ageing and Health,
Newcastle University,
Newcastle-upon-Tyne, U.K.
e-mail: e.k.perry@ncl.ac.uk

Adrian Pichurko

Marcus Raichle

Mallinckrodt Institute of Radiology,
Washington University School of Medicine,
St. Louis, MO, U.S.A.
e-mail: marc@npg.wustl.edu

Dean Radin

Institute of Noetic Sciences,
Petaluma, CA, U.S.A.
e-mail: dean@noetic.org

Antti Revonsuo

School of Humanities and Informatics,
University of Skovde, Sweden.
Centre for Cognitive Neuroscience,
University of Turku, Finland.
e-mail: revonsuo@utu.fi

Ashish Ranpura

Centre for Advanced Biomedical Imaging,
University College London,
London, U.K.
e-mail: ash@ranpura.com

Naren P. Rao

Department of Psychiatry, National Institute
of Mental Health and Neuroscience,
Bangalore, India.
e-mail: docnaren@gmail.com

Edmund T. Rolls

Oxford Centre for Computational
Neuroscience,
Oxford, U.K.
E-mail: Edmund.Rolls@oxcns.org

Silvio Scarone

Department of Medicine, Surgery
and Dentistry, University of Milan,
Milan, Italy.
e-mail: silvio.scarone@unimi.it

William Sedley

Institute of Neuroscience,
Newcastle University,
Newcastle-upon-Tyne, U.K.
e-mail: willsedley@gmail.com

Michael Sharwood Smith

Department of Languages and Intercultural
Studies, Heriott Watt University,
Edinburgh, U.K.
e-mail: msharwood@blueyonder.co.uk

Signy Sheldon

Department of Psychology,
University of Toronto, Canada.
e-mail: signy.sheldon@utoronto.ca

Wolf Singer

Department of Neurophysiology,
Max-Planck Institute for Brain Research &
Frankfurt Institute for Advanced Studies,
Goethe University,
Frankfurt, Germany.
e-mail: singer@mphi-frankfurt.mpg.de

Mark Solms

Department of Psychology,
University of Cape Town,
South Africa.
e-mail: mark.solms@uct.ac.za

John Truscott

Department of Foreign Languages
and Literature,
National Tsing Hua University,
Hsincha 30012, Taiwan.
e-mail: truscott@mx.nthu.edu.tw

Katja Valli

Centre for Cognitive Neuroscience,
University of Turku, Finland.
School of Humanities and Informatics,
University of Skovde, Sweden.
e-mail: katval@utu.fi

Kai Vogeley

Department of Psychiatry,
University of Cologne, Germany.
e-mail: kai.vogeley@uk-koeln.de

Ursula Voss

JW Goethe-Universitat Frankfurt,
Bonn, Germany.
e-mail: u.voss@uni-bonn.de

Prologue

Synopsis

We believe this is a unique book on consciousness. It is a fascinating cornucopia of new ideas on the subject, based on the fundamentals of neurobiology, psychology, psychiatry and therapy that extends the boundaries of current concepts of consciousness. We hope readers, not only neuroscientists and psychologists but also professionals from other quarters of the academic world with a general interest in exploring consciousness, will find this eclectic mix as stimulating and challenging as we do.

Aims

When the Editor of this prestigious series, Maxim Stamenov, suggested a new edition of the volume published in 2002, 'The Neurochemistry of Consciousness', we found ourselves confronted with a challenge. The 1990s, and turn of the century, saw a plethora of texts on the neuroscience of consciousness following Christof Koch and Francis Crick's seminal notion that pursuing neural correlates of consciousness would unravel novel mechanisms and theories. These new books, based on brain mechanisms like neurotransmission, psychological functions like attention, and methodologies such as brain imaging, relevant as they are, have not (as might be expected) taken us nearer the central issue – how consciousness can be explained in terms of brain activities. Neither have they led to widely accepted radical paradigm shifts, such as have occurred in other areas of science like quantum physics. With time, even some die hard mechanistic neuroscientists have come to acknowledge that there may be an impasse – a gap between top-down subjective awareness and bottom up brain mechanisms as we currently understand them; the so called 'hard' problem of 'qualia', – that is not apparently bridged in current approaches.

The subject nevertheless continues to intrigue scientists and philosophers. Neuroscience, we believe, continues to play a central role in providing information on the brain that will inform thinking on the subject. Enlisting two new Editors and after much debate we decided to encourage some thinking 'outside the box'. We extended a wide net by approaching experts in the area of the neuroscience

of consciousness, seeking new concepts and approaches. We were greatly encouraged to elicit positive and enthusiastic responses to our plea for short, sharp contributions from a number of top researchers either working directly in the field of consciousness or in related areas. Although we initially suggested a focus on conscious – non-conscious interactions, many contributions have gone beyond this boundary and demonstrated the diversity of ideas that the term ‘consciousness’ can elicit.

There were various ways we could have grouped the contributions: as we had done in the previous edition into normal states (controlled or otherwise) and disease or drug induced states; bottom up – how consciousness is related to and affected by molecular/neuronal/systems, versus top down – how consciousness itself affects these brain mechanisms; academic approaches relating to theories of consciousness versus those with more pragmatic implications.

More than most books on consciousness, this volume has been structured by the ideas of individual contributors rather than any preconceived editorial plan. As ideas and concepts accumulated it seemed to us that the most logical and meaningful categorization was fourfold grouping chapters which:

1. Potentially pave the way for new research into consciousness relating to basic scientific – physiological, pharmacological or neurochemical – mechanisms underpinning conscious experience (the ‘bottom up’ approach).
2. Provide directions based on how psychological processes are involved in consciousness and related brain activity (the ‘top down’ approach).
3. Indicate how including consciousness in the equation could lead to new approaches to understanding mental disorders, such as schizophrenia, depression, dementia, psychosis in general (in particular hallucinations) and addiction, including, amongst therapeutic approaches, an understanding of the power of the placebo.
4. More provocatively, but still based on scientific evidence, explore ways that consciousness can be considered beyond conventional boundaries, including some which indicate the potential for radical new thinking or possible ‘quantum leaps’ in neuroscientific theory.

These divisions are not absolute and one of the book’s strengths is that each chapter stands alone and can, therefore be read in any order, over any period of time.

Below each Editor has highlighted what they consider to be the main points of interest for each of the contributions in the book. We hope this overview will guide the reader to chapters related to their particular field of interests, but also that it will entice readers into areas outside their fields of expertise. This has certainly been our journey as Editors and we hope that it is one that readers will also undertake and enjoy as much as we have done.

Neuronal mechanisms

Slow cortical potential hypothesis of consciousness. Biyu He and Marcus Raichle postulate that the slow cortical potential recorded from the surface of the brain is a key neural substrate that could facilitate integration across wide cortical areas – a prerequisite for conscious awareness. In view of the link between the slow cortical potential and the fMRI signal, their hypothesis can be tested empirically.

Distinct characteristics of conscious experience met by large scale neuronal synchronization. Lucia Melloni and Wolf Singer propose that synchronization of distributed neuronal activity patterns meets most requirements for the neuronal mechanisms supporting consciousness. The challenge of identifying neuronal correlates of consciousness is distinguishing between the processes that lead to conscious experience and those that follow once contents have become conscious. Criteria for this distinction are outlined, separating the contribution of each subsystem and characterizing how they interact; providing a real alternative to the ‘ever lurking homunculus’.

Cellular components of gamma oscillations. Fiona LeBeau explores the role of key neuronal elements in the generation of cortical gamma frequency oscillations. She discusses how new techniques for labelling and activating in isolation specific sub-classes of neurons (optogenetics) can be used to explore the cellular components that underlie the generation of a brain activity that may be specifically linked to aspects of consciousness.

Dopaminergic decision making. Anthony Grace discusses how changes in dopamine levels in the limbic system (prefrontal cortex, hippocampus and nucleus accumbens) guide behavioural responses and modulate decision-making. He proposes that dopamine can serve to focus conscious attention to a particular task and can flexibly alter behaviour to achieve goals.

Endocannabinoid system and undercurrents of consciousness. Heather Ashton stresses that unconscious processes set the ‘tone’ for all conscious moods, thoughts and feelings. These unconscious processes are mediated by modulatory neurotransmitter systems, of which the endocannabinoid system is one of the most powerful. This system modulates almost all vital functions including those clearly involved in consciousness (cognition, mood and others) and many that do not normally reach consciousness (motor control, endocrine activity and others).

Disconnecting consciousness in general anaesthesia. George Mashour and colleagues explore molecular and cognitive mechanisms of general anaesthesia and their relevance to the science of consciousness. They suggest that general

anaesthetics serve as an important but still relatively under used tool to explore transitions from conscious to unconscious processing, as these drugs rapidly induce a complete and reversible cessation of consciousness. With vast numbers (over four million patients each year in North America alone) undergoing general anaesthesia, the highly controlled and monitored setting of the operating room could become the ultimate consciousness laboratory.

Neural mechanisms of mental time travel underpin the continuity of consciousness across time. Michael Hasselmo uses the term neural time travel to describe the processes of drawing on distributed cortical perceptual processes for detecting the state of the self along multiple dimensions. These include spatial location, head direction, speed, temporal duration, and egocentric relationships to items. He provides evidence that one particular set of cerebral circuitry, mediated by cholinergic muscarinic receptor mechanisms, activates neuronal properties involved in seamless internal representation of self that is based on memory.

Psychological processes

Explicit and implicit memory share underlying processes. Signy Sheldon and Morris Moscovitch remind us of the commonly held assumption that consciousness is a defining feature that distinguishes explicit memory (with conscious awareness) from implicit memory (without conscious awareness). Although early studies support this notion, recent evidence suggests that conscious and non-conscious memory systems may share crucial underlying processes. They propose that one locus of interaction between some types of explicit and implicit memory may be the non-conscious processes associated with recollection or detailed remembering, that are mediated by the hippocampus.

More than one type of non-conscious processing. Stan Franklin and Bernard Baars suggest that unconscious processes come in two varieties – the preconscious, whose contents may become conscious, and the never-conscious, whose contents may not. They enlist their Global Workspace Theory and a related model to catalogue never-conscious and preconscious processes. They suggest that the functional distinction between never-conscious and preconscious processes derives from one of the major purposes of the consciousness mechanism – to select the most salient portion of the current situation to which to attend and to broadcast this globally, in order to choose the best next action.

Gambling as a measure of awareness. Colin Clifford and colleagues remind us that most of the processing of incoming sensory information is not accessible

to consciousness. These authors are concerned with methodological issues in exploring the limits of unconscious processing and evaluate the intriguing tool of post-decision wagering as a candidate method for measuring awareness.

Magical methodology. Gustav Kuhn shows us how magicians have developed powerful techniques to manipulate our perception and awareness. Many of these techniques share similarities with phenomena typically investigated by psychologists and neuroscientists. His novel approach to the study of consciousness is to utilize the magician's expertise to complement more traditional experimental laboratory based research. This approach offers new and exciting insights into wide areas of consciousness, such as attention, visual awareness and how top-down processes modulate perception.

The advantage of a noisy brain. Edmund Rolls focuses on decision making and the advantages of noise caused by randomness in the spiking times of neurons in the brain. Noise results in the brain operating effectively as a non-deterministic system, which has implications for free will. It also results in decisions being taken probabilistically between the reasoning system and the implicit reward system. If free will describes the operation of the reasoning system, consciousness is then a property of a reasoning system that must use higher order syntactic thoughts to correct its first order thoughts. Decision making in an implicit system potentially involves confabulating a reason for the decision with the feeling of being free an illusion.

Social consciousness as a key function of the default network. Kai Vogeley describes the neural correlates of the reciprocal nature of social interaction using fMRI. Self-consciousness, as awareness of ones own mental states, and social consciousness, as knowledge of the mind of others, overlap with the default mode of brain function. A key biological function of this default network, irrespective of task, may be social consciousness. Humans thus have a built in disposition for social cognition that is reflected in this neural default mode.

Linguistic processing at two levels. Mike Sharwood Smith and John Truscott point out that knowledge, use and acquisition of language is largely an unconscious process for a child acquiring its first language. However, it leaves conscious footprints in the form of a 'voice in the head'. Learning a second language in later years, the mature learner can and often does develop a conscious understanding of what is being acquired. This understanding springs from a knowledge that is separate from the unconscious and inaccessible system developed within the language module. It is an open question how much that separate knowledge and the voice in the head that it engenders actually enhances or hinders development and on-line performance.

Dreaming as key to understanding consciousness. Antti Revonsuo and Katja Valli suggest that dreaming as a state of subjective awareness, essentially independent of input and output, best lends itself to the challenge of bridging the top-down / bottom-up divide. They go so far as to indicate that if the internally generated state of dreaming is not accounted for by any neuroscientific theory of consciousness that theory is lacking.

Lucid dreaming and the bimodality of consciousness. Allan Hobson and Ursula Voss argue that, in the context of the widely accepted division between primary and secondary or reflective consciousness, dreaming is dominated by the primary type. Intriguingly, lucid dreaming, during which rational thought and volition emerge in a state of sleep dominated by primary consciousness, provides a new investigative tool. Exploring dreaming is proposed as a cornerstone of higher level consciousness research.

Psychopathologies and therapies

Why depression feels bad. Mark Solms and Jaak Panksepp highlight the urgent need to restore conscious experience to psychiatric practice. They provide evidence for the failure of bottom-up antidepressant drugs or theories of serotonin deficits to deal with the problem of depression. Their idea of depression as a kind of bereavement based on loss of self connectedness and separation is, they suggest, a pathological extension of a natural survival mechanism. Accordingly continuing to ignore such conscious experience is, they suggest, likely to be perilous in clinical practice

Consciousness abhors a vacuum. Daniel Collerton considers the question of what kind of consciousness is experienced by people with dementia and suggests they have not diminished, but different conscious awareness. Resulting from brain pathologies which impair perception and other aspects of cognition, gaps in normal conscious awareness are not left unoccupied but are 'filled in' resulting in, for example, hallucinations or delusions. This view encourages empathy in those interacting with people with dementia that could impact on care strategies.

Consciousness as the spin-off and schizophrenia as the price paid for language in man. Tim Crow puts forward a novel and thought provoking theory of schizophrenia based on the asymmetry (torque) of the human brain as the foundation of the faculty of language. Focusing on the symptoms of hearing voices, the experience of thoughts as not one's own, and incoherent speech as core symptoms of schizophrenia, he follows three new lines of argument. First, this torque is

the feature that defines the human brain as four chambered by comparison with the two chambers of the generalized mammalian brain. Secondly, by separating thought from speech production in the frontal lobes, and meaning from speech perception in occipito-parieto-temporal association cortex, torque confers on our species the capacity for language. Thirdly, the phenomena of psychosis can be seen as 'leakage' from one to another of the four quadrants of association cortex.

Hippocampal seat of consciousness and hallucinations in schizophrenia. Ralf-Peter Behrendt argues that a hippocampal auto-association network (CA3) pinpoints a spatiotemporal and emotional context for the purpose of event memory formation. Representing a continuous flow of complex symbols, conscious experience may be irrelevant to the workings of the brain, although neural activity underlying event memory formation influences behaviour (via output to medial prefrontal cortex, ventral striatum and lateral septum). Excessive pyramidal cell activity in the CA3 area, due to deficient inhibition (by GABAergic basket interneurons), leading to event memory formation unrestrained by input from areas such as entorhinal, may be a mechanism for the generation of altered consciousness such as hallucinations in schizophrenia.

Perspectives from the Charles Bonnet Syndrome. Dominic ffytche examines evidence derived from visual hallucinations in the context of eye disease – the Charles Bonnet Syndrome – to examine the nature of the visual unconscious. Forcing us to reconsider the nature of the unconscious, this hidden system of processing underlies our apparently seamless conscious experience of the world, with many of its complex functions yet to be recognised by visual science.

Believing is hearing is believing: the reciprocal nature of consciousness. Will Sedley provides an analysis of the neural correlates of auditory hallucinations including tinnitus, musical and verbal hallucinations, misperceptions and imagery. All types of auditory consciousness involve auditory cortex activation which can often be measured as synchronised gamma band oscillations, but disruptions in connectivity, such as cochlear damage, can lead to positive feedback cycles that cause and enhance hallucinations.

Dreaming as a model of psychosis. Armando D'Agostino, Ivan Limosani and Silvio Scarone investigate the long neglected dream state of consciousness in psychiatry. They argue that the similarity between normal dreaming and states of psychosis provides a novel fertile area for psychopathology and consciousness research. In particular, they suggest that lucid dreaming, whereby there is awareness of the dream state during sleep, whether induced by psychological or pharmacological means, could be a new tool for investigation and therapy in psychiatry.

If only drug users were aware of why they choose to use. Andrew Parrrott presents new evidence based on reports by users of ecstasy/MDMA that their choices are governed by largely non-conscious optimistic beliefs focused on desired aims rather than actual consequences, with low levels of or little attention to, conscious awareness of all the issues. He describes how users become aware of the adverse effects of MDMA and make the conscious decision to use less frequently and quit permanently.

Brain body interactions in placebo responses. Fabrizio Benedetti presents a host of fascinating data on how placebo responses, involving a range of cues, depend on both conscious and non conscious processes, and how these can mimic therapeutic drug effects even inducing cellular and molecular changes in the patient's brain (e.g. receptor function or neuronal circuitry). He argues that this evidence provides the basis for the efficacy of cognitive and other psychological therapies in terms of neurobiological mechanisms. This raises the intriguing question of how we might deliberately self modify our mindset at both conscious and non conscious levels to promote healing processes.

Expanding boundaries

Paradox of creativity. Ashish Ranpura and Mark Lythgoe describe a neurobiological theory that explains the process of creativity – possibly the most impressive facet of secondary consciousness. The authors explore how three paradoxes of creativity, the paradox of ego, of focus, and of quality are defining features of this quintessentially human experience. The frontal lobes appear to play a critical role in creativity and the authors draw upon a Darwinian selection theory that allows conscious creativity to emerge from a series of unconscious mental activities.

How research on meditation contributes to the neuroscience of consciousness. Antoine Lutz examines initial findings of neuroscientific research on meditation and reveals how these identify new horizons of further inquiry in consciousness research. While such studies of contemplative practices are still in their infancy, early findings promise to contribute in three key areas. These include: neuroplasticity – physiological and psychological indices of short and long terms responses of the brain circuits that underlie complex mental functions associated with specific types of meditation techniques; mind body interactions – revealing mechanisms by which such training may exert beneficial effects on physical health; and subjectivity – well developed introspective skills of practitioners potentially shedding new light on the neural counterpart of subjectivity.

Exploring consciousness based on self induced altered stated techniques. Bangalore Gangadhar and Naren Rao provide a fascinating perspective on methods used to alter personal conscious experience using a variety of procedures commonly practised in India including meditation. Together with Eastern concepts of the nature (levels or states) of consciousness which diverge markedly from Western ideas, for example belief in a cosmic consciousness as a source of all conscious experience, he provides food for thought on potential paradigm shifts in theories of consciousness. At the same time, how physiological and brain imaging markers are affected as a result of meditation provides neuroscience with down to earth tools for investigating an area hitherto regarded by most neuroscientists as esoteric at best.

Models of conscious and non-conscious perception may need radical revision in unexpected ways. Dean Radin provides evidence from anecdotal to controlled scientific experimentation (published in refereed journals) that telepathy is a genuine phenomenon. He raises the question of how to account for such conscious interactions beyond the 'common' senses and what kind of new neuroscientific theories might have to be generated. An intriguing theory of 'entanglement' based on principles of quantum physics is discussed as a possible explanation of such 'non local' phenomena – food for thought for neuroscientists looking beyond continuing shortcomings of reductionism to explain conscious experience.

Plants of the gods and shamanic journeys. Elaine Perry and Valerie Laws suggest that knowledge and experience of shamanic practitioners may have more to offer consciousness studies than psychedelic plant chemicals. Among agents which alter the boundary between conscious and non-conscious cognition, the ritualistic use of plant species provides an example of long-standing empirical knowledge subsequently verified by scientific (chemical, pharmacological and psychological) evidence. Based on such an impressive 'track record' for obtaining validated information about consciousness, further investigation of shamanic experiences including reports of other 'dimensions' of consciousness are warranted. Evidence obtained in such areas, rooted firmly in scientific methodologies (though no doubt incorporating factors not normally part of scientific methodology such as mind training, belief, past experience and empathy), could potentially contribute material for radical new scientific theories of consciousness,

Conclusions

This book clearly reflects the individuality of its contributors more than most. In seeking novelty, it contains new 'streams of consciousness' on the subject of neuroscience and consciousness that we think provide new directions for future research.

We would hardly expect readers, any more than the Editors as it transpired, to agree with all the arguments raised within this book. A unitary definition of consciousness was neither sought nor provided. In originally setting an interface between conscious non conscious processes as a target focus, we may have been overly optimistic that agreement on any such division could be reached. Categorizing distinct levels of consciousness or considering continuity are clearly far from being resolved. Contributors have incorporated a range of definitions or states of consciousness in addition to or other than the conscious non conscious divide: Gerald Edelman's division between primary or first order and secondary or reflective consciousness; divisions within non conscious processing itself; self consciousness; dreaming as a neglected but core state for investigating consciousness; and Eastern concepts of multiple states that challenge the Western mind.

Drawing together what we consider the most stimulating and challenging ideas into emergent themes for neuroscience based consciousness research, key questions emerge:

Are boundaries between conscious and non conscious processing less obviously demarcated than previously considered, with multiple levels or states or perhaps a boundless continuum?

How much of cognitive or perceptual conscious awareness is based on illusion versus reality?

How many other non-conventional activities and investigative tools (like magic and gambling) might be incorporated into research into consciousness?

Is it time to reinstate dream research into brain research as a model for investigating the essence of conscious awareness in normal and pathological states?

Which if any, or all, of candidate electrophysiological measures is most likely to reflect the mechanisms which are specifically involved in conscious awareness?

Do individual signalling (e.g. neurotransmitter) systems govern specific aspects of consciousness and, if awareness is but a minute portions of pre/unconscious processing, is any one system primarily in control of the transition?

Why is general anaesthesia not a mainstream tool in more basic consciousness studies?

Can psychiatric practice thrive or survive without attending to the state and nature of consciousness in individual patients?

Can pathologies associated with common disorders of the human mind be analysed more in relation to consciousness and so inform basic mechanisms of consciousness?

If the placebo effect is so robust and associated with the same mechanistic changes in brain and body as those induced by drug treatment, why is it not the subject of more therapeutic research?

Is it time for neuroscientists to seriously consider scientific evidence for so called anomalous phenomena such as telepathy, and in this context contemplate paradigm shifts in theories of consciousness?

Should more attention be given to concepts of time honoured mystics about consciousness which could be relevant to neuroscientific enquiry?

In the process of gathering contributions for this book it became clear that a second volume could easily have been generated, as there are many topics we have not had sufficient space to cover in this volume. For example, consciousness in the vegetative state springs to mind as an omission that reminds us of the need for new tools to assess individual subjective states of conscious awareness. New neurophysiological findings of a novel type of high speed cohesion between distant groups of neurons remind us how exciting it is when a new and unexpected observation emerges at the level of observational science to inspire new thinking about neural network integration relevant to consciousness. Default mechanisms that include a kind of continual 'day dreaming' as a fall back position in the awake state remind us of the essentially and continuously creative nature of the conscious stream. And brain body interactions referred to in passing in some chapters may cause us to consider consciousness as a more global phenomenon.

Acknowledgements

The Editors gratefully acknowledge: Maxim Stamenov for not only for suggesting this new volume but also for his guidance and wisdom in the shaping of the book; Valerie Laws and Susan Aldworth, currently Writer and Artist in Residence respectively at Newcastle University, for the poetic and artistic inclusions; Brian Moore for suggesting contributors; Peter Whitehouse for formulating ideas for the prologue; Peter Fenwick for guidance on 'beyond the brain' issues; Peter Russell and Ervin Laszlo for agreeing their article on the primacy of consciousness could be included (and we refer readers interested in this con-

cept, relevant to some of the chapters in the last section to the relevant website [<http://www.redicecreations.com/specialreports/2006/02feb/primacyconsciousness.html>]); our respective Heads of Department for supporting our work on the book, and to the Publishers John Benjamins.

Seeing myself think

Last night I saw myself think;
I caught myself out. Just
As I went to sleep
I saw the ideas, a sea
Of winking sparks like mackerel
Under the moon or the flutter
Of cameras flashing
In a dark auditorium; or
The bioluminescence
Of deep sea creatures signalling
In the black cold. Each thought
Was a light leaping a synapse
Where pairs of nerve cells
Reached out to each other, their touch
Kindling. Whole galaxies
Of consciousness glittered before
My eyes.

And I saw myself thinking, our sun
Might be a spark of thought
In an enormous mind
Which sees its thoughts flash
In the birth and death of stars.

Valerie Laws
(formerly published in *Quantum Sheep*,
Peterloo Poets)

SECTION I

Neuronal mechanisms



Birth of a Thought 2. Susan Aldworth (2007) etching and aquatint, 35 × 50 cm

The slow cortical potential hypothesis on consciousness

Biyu J. He and Marcus E. Raichle

We propose a neurophysiological hypothesis on the emergence of consciousness, which postulates that the slow cortical potential (SCP) recorded from the surface of the brain provides an index of the activities of superficial-layer pyramidal neurons that directly contribute to the emergence of conscious awareness. This hypothesis is supported by existing data from manipulations of conscious awareness in normal human subjects and by data from altered states of consciousness such as general anesthesia and recovery from vegetative states; it further makes experimentally testable predictions. Given a relationship between the SCP and the functional magnetic resonance imaging (fMRI) signal that has now been revealed, this hypothesis also provides a potential bridge between existing neuroimaging and electrophysiological studies on consciousness.

Since its introduction in the early 1990's, fMRI has become the most widely used tool in cognitive neuroscience and has produced a formidable array of brain maps depicting both localization (as in traditional activation studies) and integration (as in more recent functional connectivity studies) of brain activity. As the fMRI signal measures directly blood oxygenation and only indirectly neuronal activity, an important need for understanding the neural events contributing to the fMRI signal has been widely recognized. Such a need is further stressed by the inconsistencies between a number of human fMRI and monkey unit physiological studies employing the same tasks (Logothetis, 2008). Responding to this need, a number of studies have compared the fMRI signal or its close relatives (including tissue oxygenation, blood flow, optical intrinsic signals) with simultaneously recorded electrophysiological signals. The convergent results from these studies suggest that the fMRI signal is contributed predominantly by synaptic activity representing inputs and local processing in an area as measured by local field potentials (LFP). The spiking activity, though often correlated with both the LFP and the fMRI signal, can be dissociated from the latter two in a number of conditions

including adaptation, drug modulation, manipulations of excitatory and inhibitory inputs, and a spatial separation between input and output activity (for reviews see (He and Raichle, 2009; Logothetis, 2008)).

Whereas multiple frequency ranges of the LFP (e.g., 5–30 Hz, 20–60 Hz, ~25–90 Hz, see references in (He and Raichle, 2009)) have been correlated with the fMRI signal in different conditions, all of these studies have only assessed power modulations of the LFP because only the power of these frequency ranges has a comparable temporal scale to that of the fMRI signal (<0.5 Hz). Recently, it was found that the raw fluctuations of the low-frequency end of field potentials (<4 Hz), with a temporal scale overlapping that of the fMRI signal, also correlates with the fMRI signal (He and Raichle, 2009; He et al., 2008; Khader et al., 2008; Nagai et al., 2004). This signal, termed the “slow cortical potential” (SCP) by us and others (Birbaumer et al., 1990; He and Raichle, 2009; Khader et al., 2008), appears optimally positioned for carrying out large-scale information integration in the brain. Because conscious awareness (see Glossary) is always experienced as a unitary and undivided whole, rather than components that are experienced independently (in the case of split-brain patients, there are two conscious experiences each located in a separate brain hemisphere, rather than one single conscious experience), segregated information processing in the brain cannot contribute directly to conscious awareness. We propose that the neuronal activities reflected by the SCP may contribute to the emergence of consciousness and review existing empirical evidence supporting this idea. As the current hypothesis is based on a well-defined, well-characterized physiological process, it is entirely amenable to empirical testing.

The SCP and the fMRI signal

The SCP is the slow end (mainly <1 Hz, can extend up to ~ 4 Hz) of brain field potentials that can be recorded using either depth or surface electrodes. A negative shift in surface-recorded SCP indexes increased cortical excitability (for detailed physiology see the following section). Since the SCP frequency range is subject to artifacts due to sweating (in scalp-electroencephalography (EEG) recordings), movement and electrode drift (if polarizable electrodes are used), it has been eliminated in most animal physiology as well as human EEG studies by online high-pass filtering. This is unfortunate because, as was recognized in the 1970’s, “If DC [i.e., direct-current] recording is used, virtually every stimulus-bound cortical activity is seen to be accompanied by a change in cortical steady potentials” (Goldring, 1974).

Recently a number of studies have revealed a close relationship between the SCP and the fMRI signal. Nagai et al. (2004) demonstrated covariation of simultaneously recorded SCP and fMRI signals during task stimulation using simultaneously recorded EEG and fMRI (Nagai et al., 2004). These authors found a trial-by-trial correlation between the amplitude of a negative SCP response indexing expectancy (“contingent negative variation”, CNV) and the fMRI signal amplitude in anterior cingulate cortex. The anterior cingulate has previously been determined as a generator region of CNV (Nagai et al., 2004). Jones et al. (2007) showed that spontaneously fluctuating total haemoglobin concentration (a signal tightly linked to the fMRI signal) and low-pass filtered LFP (i.e., depth recorded SCP) are temporally correlated (Jones et al., 2007). Our own data using invasive EEG (i.e., electrocorticography, ECoG) and fMRI in neurosurgical patients showed that large-scale (2–10 cm on cortical surface) correlation patterns in the spontaneous SCP and fMRI signals were similar (He et al., 2008). This finding has since been extended to inter-hemispheric correlations as well (unpublished data). Beyond these approaches, there is an extensive literature showing similar modulation patterns of the SCP and the fMRI signal in a wide range of cognitive tasks (for reviews see (Birbaumer et al., 1990; He and Raichle, 2009; Khader et al., 2008; Rosler et al., 1997)). Hence, convergent results suggest that the SCP has a close correspondence to the fMRI signal in different experimental conditions. Next, we consider the physiological mechanisms underlying the SCP.

The physiological basis of the SCP

Simultaneous recordings of surface potentials, field potentials in different cortical layers, and intracellular membrane potentials have clearly demonstrated that synaptic activities at apical dendrites in superficial layers are the main factor contributing to the SCP. Specifically, long-lasting excitatory postsynaptic potentials (EPSPs) at these apical dendrites underlie negative-going surface-recorded SCPs (Birbaumer et al., 1990; Goldring, 1974; Mitzdorf, 1985). As an example, we consider the effect of visual stimulation in V1 – a standard model for the investigation of fMRI-electrophysiology correspondence. Specific thalamic inputs terminate first on the soma of layer-IV stellate cells and layer-III pyramidal cells, and then follow one of two pathways to depolarize the apical dendrites of superficial- or deep-layer pyramidal cells (Mitzdorf, 1985) (Figure 1a).

Given the geometry of cortical fields, the earlier processes – excitations of pyramidal cells at their soma, produce positive-going surface potentials (Figure 1b i). The later processes, excitations of pyramidal cells at their apical dendrites, produce surface negative potentials (Figure 1b ii and iii). However, EPSPs at apical

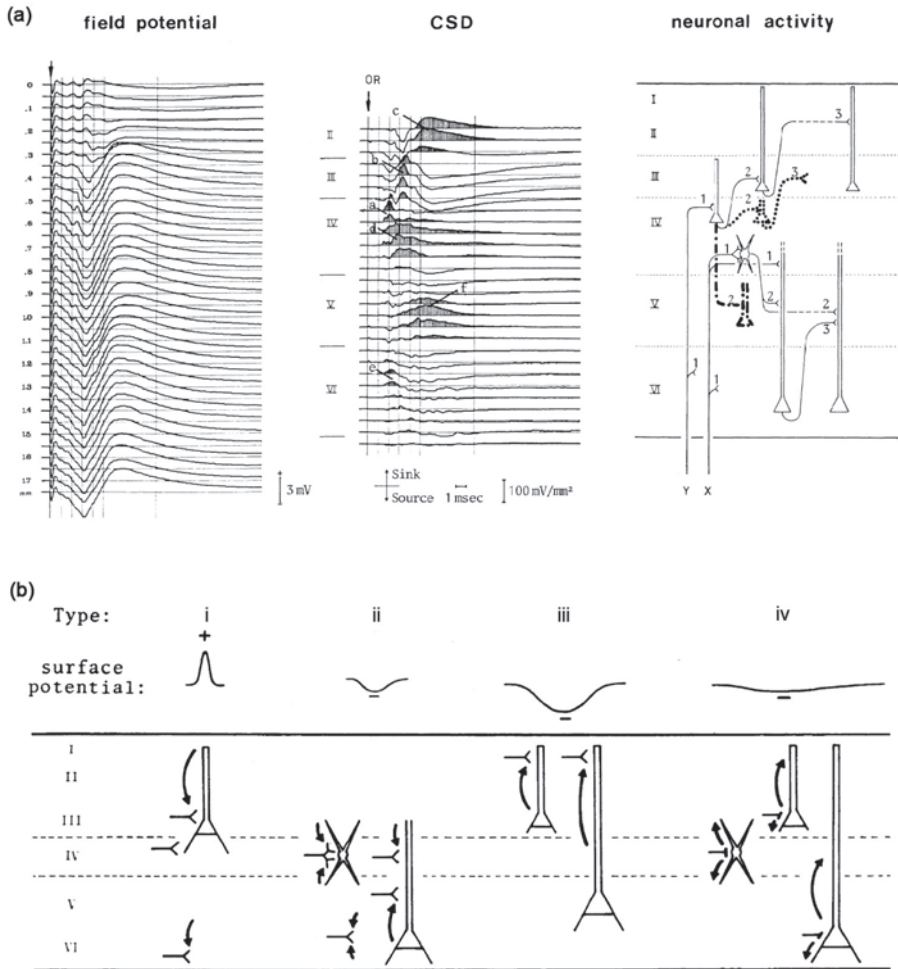


Figure 1. The physiological basis of the SCP. (a) *Left:* field potential in primary visual cortex of the cat evoked by electrical stimulation of optical radiation. Each trace is the average of 20 responses. Distance between adjacent recordings is 50 μm . *Middle:* current source-density (CSD) distribution obtained from the potential profile on the left. Sinks, corresponding to active EPSPs, are shaded. Cortical laminae are indicated. Sinks a, b, and c reflect mono-, di-, and trisynaptic Y-type activity as shown in the right panel; sinks d and f reflect mono-, di-, and trisynaptic X-type activity shown in the right panel; sink e reflects Y-type and X-type monosynaptic activity. Sinks a, b, and e contribute to type *i* activity in (b); sinks d and f contribute to type *ii* activity in (b); sink c contributes to type *iii* activity in (b). *Right:* schematic diagram of successive intracortical excitatory relay stations as well as cell types involved. Long-range feedback connections and nonspecific thalamic inputs are not depicted. Numbers 1, 2, 3 denotes mono-, di-, and tri-synaptic activity respectively. (b) Schematic diagram of

dendrites of deep layer pyramidal cells create closed fields and thus have rather small influence on surface potentials (Figure 1b ii). By contrast, depolarizations of superficial layer apical dendrites contribute greatly to negative SCPs (Figure 1b iii). The contribution of inhibitory interneurons to SCP or field potentials in general is also small because of the low amplitude of membrane current flow during inhibitory activity and a lack of laminar specificity (Figure 1b iv) (Birbaumer et al., 1990; Mitzdorf, 1985). In summary, the later component of sensory evoked potentials in EEG or ECoG recordings – a negative slow-potential shift – is primarily due to long-lasting depolarizations of superficial layer apical dendrites.

Other than activations by specific thalamic inputs described above, the superficial layers are also where long-range intracortical (i.e., intra-areal) and cortico-cortical (i.e., inter-areal) connections preferentially terminate (Braitenberg and Schuz, 1998; Douglas and Martin, 2004; Mitzdorf, 1985). First, only in superficial layers do pyramidal cells make extensive horizontal arborizations (Douglas and Martin, 2004). Thus, EPSPs in superficial layers spread over a considerable spatial extent and manifest themselves as “depolarization fields” (~several mm²) in optical imaging recordings (Roland et al., 2006). Second, long-range inter-areal feedback connections also terminate mainly in superficial layers. Hence, it is not surprising that the SCP and the correlated fMRI signal reveal large-scale brain networks in their spontaneous fluctuations (He et al., 2008). Moreover, superficial-layer apical dendrites are also the main target of nonspecific thalamic inputs that originate from “matrix cells” spread throughout the thalamus (Jones, 1998). Interestingly, the reticular thalamic nucleus, which the nonspecific thalamocortical projections must pass through, exerts a low-pass filter influence that may facilitate the emergence of slow activity (He and Raichle, 2009). In summary, long-range intracortical and feedback cortico-cortical connections, as well as the nonspecific thalamic inputs, all contribute directly and significantly to the SCP.

4 main types of cortical activities and their reflection in surface potential (recorded by ECoG or EEG). *i*: Depolarization of pyramidal cells at their deeper extremities, which generates a surface-positive potential deflection. *ii*: Depolarization of deep-layer pyramidal cells at their apical dendrites or stellate cells. This type of activity generates a sink in the middle layers and a surface-negative potential deflection. But because of the closed-field arrangement of CSD components, its contribution to surface potential is rather small. *iii*: Depolarization of superficial layer pyramidal cells at their apical dendrites. This is the main contributor to long-lasting surface-negative potentials. This type of activity involves long-distance connections and depends greatly on the general state of cortical excitability. *iv*: Inhibitory activity does not usually cause significant CSD contributions, because of the low amplitudes of membrane currents involved and a general lack of lamina specificity. Adapted with permission from Mitzdorf (Mitzdorf, 1985).

Given that negative SCPs index increased cortical excitability, it should not come as a surprise that during the negative shift of spontaneous SCP fluctuations there are increased multi-unit activity, increased higher-frequency field potentials, and better behavioral performance (see references in (Birbaumer et al., 1990; He and Raichle, 2009; He et al., 2008)). The recently acclaimed phase-coding in the delta frequency range (e.g., (Lakatos et al., 2008)) is likely of the same origin as information carried in the SCP phase. Moreover, the P300 evoked potential, a proposed correlate of conscious neuronal activity (Dehaene et al., 2006), might reflect inhibition of cortical pyramidal neurons and constitutes the flip side of negative SCP shifts (Deecke and Lang, 1988). Since the SCP modulates power of higher-frequency activities, it may be a more fundamental correlate of the fMRI signal than LFP power is, as implicated in a previous study (He et al., 2008).

As mentioned previously, the SCP is one important and substantial contributor to the fMRI signal. In addition to advancing our understanding of the fMRI signal and bridging the neuroimaging and neurophysiology fields, this observation is also of particular interest in the study of consciousness. For example, fMRI experiments and single-unit recordings often show discordant results during manipulations of consciousness; this disagreement has been most dramatic in V1 (Logothetis, 2008; Tononi and Koch, 2008). These puzzling results are at least partially illuminated when we bring the SCP and its underlying physiology into the picture. In what follows we discuss these data and further propose a specific hypothesis on the involvement of the SCP in engendering conscious awareness.

The SCP and consciousness – A neurophysiological hypothesis of consciousness

From a theoretical perspective, information has to be integrated to contribute to conscious awareness, for conscious experience is always a unitary and undivided whole. We suggest that the superficial-layer neuronal activities indexed by the SCP may be an optimal neural substrate to carry such information integration across wide cortical areas because (1) its slow time scale allows synchronization across long distance despite axonal conduction delays (He et al., 2008; von Stein and Sarnthein, 2000); (2) long-range intracortical and corticocortical connections terminate preferentially in superficial layers and thus contribute significantly to the SCP. Furthermore, for each patch of superficial-layer pyramidal neurons (for definition of “patch”, see (Douglas and Martin, 2004)), corresponding deep layer neurons could provide additional information through specialized local processing. These local deep-layer loops might constitute a neural substrate for unconscious processes that can affect and be affected by conscious experience. A rough schematic depicting our hypothesis is shown in Figure 2a.

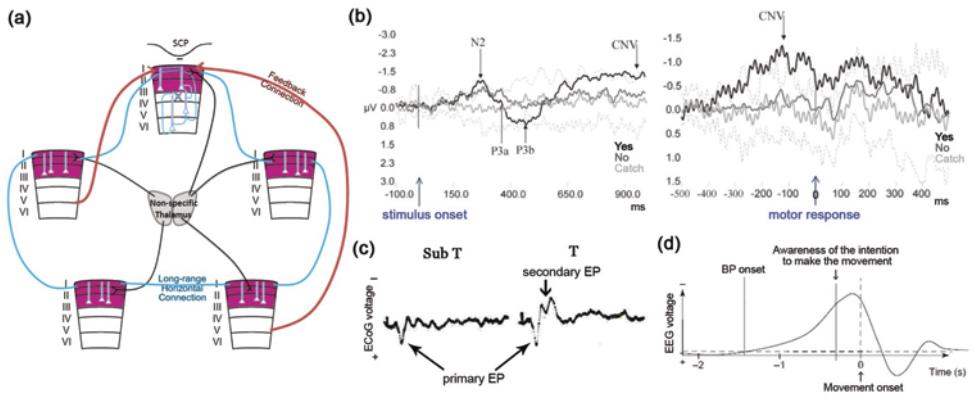


Figure 2. SCP and consciousness. (a) Schematic illustration of our hypothesis. Superficial layers of the cerebral cortex (shown in purple) are the only layers containing extensive long-range horizontal connections (thick blue lines); they are also the main target for nonspecific thalamo-cortical inputs (black lines) as well as long-range inter-areal feedback connections (thick red arrows). We propose that long-lasting synaptic activities in superficial layers, manifesting as SCPs in surface recording or low-frequency current source density (CSD) activity in superficial layers, carry large-scale information integration in the brain and contribute directly to conscious awareness. Neuronal circuits in deep layers (thin blue lines) provide specialized local processing that assist superficial-layer computations and send output to subcortical structures. Two specific predictions made by this hypothesis are provided in Conclusions. (b) Subjects performed a target detection task in which a visual grating stimulus at threshold was briefly presented. Following a variable delay, the subject was prompted by an auditory cue to press one of two buttons to indicate whether they saw the stimulus. A small percentage of catch trials in which no grating was presented were randomly interleaved. EEG potential from the left parietal electrode (P3, using Laplacian derivation, which emphasizes local vertical currents underneath the electrode) was averaged around the onset of grating stimulus (*left panel*), or around the motor response (*right panel*). The evoked potentials for “Yes, I saw”, “No, I did not see”, and catch trials are shown in black, dark grey and light grey respectively. The inter-subject s.d. for catch trials are shown as dotted lines. A negative slow potential builds up between stimulus onset and motor response during “Yes” trials but not catch trials nor the trials during which the stimulus was present but not perceived. Adapted with permission from Pins and ffytche et al. (Pins and ffytche, 2003). (c) Average evoked-potentials (EPs) in response to single stimulus pulses at the skin, recorded from the surface of somatosensory cortex. EPs to 500 stimulus presentations were averaged for each condition. Sub T: subthreshold stimuli, none of the 500 stimuli were felt by the subject. T: threshold stimuli, subject reporting feeling some of the 500 stimuli. Each recording trace is 500 milliseconds long. Primary EP: a transient, surface-positive deflection that occurs ~ 30 ms after the stimulus, was present in both cases. Secondary EP: a later slower surface-negative component, only occurs when the stimulus was at times felt. Adapted with permission from Libet (2004). (d) The Bereitschaft potential (BP) is a negative SCP shift preceding the onset of a voluntary movement. It was shown by Libet (2004) that the onset of the BP also precedes the subject’s subjective awareness of the intention to make the movement by a few hundred milliseconds. Adapted with permission from Haggard (2008).

Interestingly, the cerebellum, generally considered nonessential for consciousness (Tononi and Koch, 2008), is notably weak in its low-frequency activity as compared to the neocortex (Bullock and Basar, 1988). The cerebellar cortex also lacks the “crowning mystery” of the neocortex – layer I, which is one major target for long-range feedback connections and nonspecific thalamic inputs in the cerebrum (Douglas and Martin, 2004; Jones, 1998). From an evolutionary perspective, the superficial layers are endowed with the largest expansion in thickness among cortical layers during mammalian evolution (DeFelipe et al., 2002). Furthermore, the nervous systems of most invertebrates (with the exception of octopus), when compared to the cerebrum of vertebrates, contain much more pronounced fast activity but notably weak slow activity, and in this sense are more similar to the vertebrate cerebellum, brain stem or spinal cord (Bullock and Basar, 1988).

In what follows we review existing empirical data supporting a functional role of the SCP in the emergence of conscious awareness.

Attention

Although attention and consciousness are distinct and dissociable phenomena (Tononi and Koch, 2008), attention clearly affects which information has better access to conscious awareness. The top-down effect of attention in early sensory cortex is largely invisible to spike recordings, but is readily seen in the fMRI signal (Logothetis, 2008). Consistent with a close correspondence between the SCP and the fMRI signal as argued here, top-down influence in V1 can be seen with measurements of the SCP using either optical imaging or field potential recordings (Lakatos et al., 2008; Roland et al., 2006). In the first case, a feedback wave of depolarization was found to traverse the superficial layers from higher-order to lower-order visual areas (Roland et al., 2006). In the second case, top-down attention was found to modulate the phase of delta-frequency activity which further modulated the power of higher frequencies (Lakatos et al., 2008). Importantly, this effect was found only in superficial layers, consistent with the physiology of the SCP and with the laminar preference of feedback connections (Douglas and Martin, 2004).

Perception

Many studies have investigated the neural correlates of conscious perception, however, only a handful presented data including the SCP, which we will focus on here. Pins and ffytche (Pins and Ffytche, 2003) presented visual stimulation

at threshold to normal subjects, so that an identical stimulus would sometimes be perceived and at other times not. In trials during which the subject perceived the stimulus, a negative slow potential builds up over parietal electrodes between the stimulus onset and the response. This slow potential was absent in trials during which the stimulus escaped conscious perception (Figure 2b). Using a visual illusion task and depth recording in V1, Leopold and colleagues (Leopold et al., 2008) showed that perceptual suppression was only associated with changes in the lowest frequencies in upper cortical layers when current source density (CSD) method (which has better localization than raw field potentials, see Figure 2a) was used. Similar to the SCP, the fMRI signal also tracks perceptual changes, whereas spiking activity was unaffected (Logothetis, 2008). Furthermore, momentary fluctuations in the spontaneous SCP have an effect on whether a stimulus at threshold is consciously perceived or not (He and Raichle, 2009). An active involvement of the SCP in conscious perception is also supported by earlier experiments in the somatosensory domain. Through a series of elegant experiments using skin stimuli, electrical stimulations applied to the subcortical pathways and the cortex itself, Libet (2004) showed that the secondary evoked potential (a long-lasting negative potential in surface-recordings), but not the primary evoked potential (a short-latency positive potential), was essential for conscious perception (Figure 2c). These findings have since received support from more recent studies (e.g., (Sergent et al., 2005)).

Volition

If consciousness is a two-sided coin, then on one side it is occupied by perception/experience; on the other side by volition/agency (Gray et al., 2007). Similar to perception, volition (i.e., voluntary actions) also has a long recognized association with the SCP. It was discovered more than 20 years ago that a negative SCP shift preceded voluntary movement and even the subjective awareness of the intention to make the movement (Libet, 2004) (Figure 2d). Though the implications of these results were highly debated by philosophers, the essential findings have been replicated numerous times and extended (Haggard, 2008). Specifically, recent fMRI experiments have determined both the brain regions underlying the intention to make a movement and those underlying the awareness of such intention (Haggard, 2008). Important in the current context, recent results showed that the outcome of a free choice can be decoded using the fMRI signal up to 6 sec before the decision outcome enters conscious awareness (Soon et al., 2008). Future experiments should illuminate whether this early fMRI signal is related to the early negative SCP shift preceding a voluntary action.

The unconscious states

One of the most dramatic experimental manipulations of consciousness comes from stimulations of intralaminar thalamic nuclei, part of the nonspecific thalamic pathways, in minimally conscious or vegetative patients (Hassler, 1978; Schiff et al., 2007) and anesthetized rats (Alkire et al., 2007). In both cases of impaired consciousness, stimulation of intralaminar thalamic nuclei significantly improved behavioral responsiveness. As mentioned previously, nonspecific thalamic afferents terminate preferentially on apical dendrites in cortical superficial layers. Therefore, stimulation of these pathways would drive negative shifts of the SCP (i.e., increased cortical excitability) and thus may restore long-range communications in superficial layers. Similarly, recovery from persistent vegetative state (PVS) was accompanied by restoration of functional connectivity between the intralaminar thalamic nuclei and prefrontal cortex (Laureys, 2005). These experiments lend important support for the current hypothesis of a relation between the SCP and conscious awareness. Further support comes from DC-recordings of auditory evoked potentials (AEP) in humans undergoing propofol anaesthesia (Fitzgerald et al., 2001). Whereas the early positive component of AEP was preserved during anesthesia, the later component – a negative shift in the SCP was abolished under anesthesia and reappeared during emergence from anaesthesia.

These results however do not suggest that the negative SCP, whenever it appears, is an index of conscious awareness. Instead, key brain regions (see Box 1) and specific inter-regional relationships may be required. Large-scale coherent structures in the spontaneous fluctuations of both the SCP and the fMRI signal have been described under unconscious states – slow-wave sleep (SWS) and deep anesthesia respectively (He et al., 2008; Vincent et al., 2007). Do these findings contradict the current hypothesis? We suggest not. These coherent fluctuations might reflect spontaneous synaptic activity constrained by anatomical connections that continue to maintain homeostasis in the brain, but which lack sufficient information content or the integration necessary for the emergence of consciousness. Indeed, in both cases, the patterns of SCP or fMRI signal coherence were weaker in the unconscious state as compared to the conscious states (He et al., 2008; Vincent et al., 2007). A decreased baseline level of cortical excitability (not assessed by temporal correlation measurements) may also contribute to loss of consciousness, as supported by aforementioned experiments in which stimulation of nonspecific thalamic pathways restored responsiveness in subjects with impaired consciousness (Alkire et al., 2007; Schiff et al., 2007).

Concluding remarks

Studies on the neural basis of the fMRI signal have focused on the LFP power. We here present evidence for another neurophysiological signal underlying the fMRI signal that has received much less attention – the slow cortical potential (SCP). The linkage between the SCP and the fMRI signal not only advances our understanding of cortical physiology but also provides a new vantage point into many experimental results. We further propose that the neuronal activities indexed by the SCP may carry large-scale information integration in the brain that contributes directly to the emergence of conscious awareness. Experiments on consciousness have seldom included the SCP, but whenever it was included, the results appear to be consistent with this hypothesis. Given that our hypothesis involves a specific physiological process, it is clearly amenable to empirical testing.

Specifically, the hypothesis put forward makes two testable predictions:

- I. Whenever there is a change in the content of conscious awareness, there should also be a concurrent change in the SCP in corresponding essential brain regions (which yet need to be determined for different forms of conscious awareness – we offer a speculation on this in Box 1).
- II. In altered states of consciousness, such as deep slow-wave sleep, under anesthesia, or in patients with severe brain injury, the spontaneous organization of the SCP that is important for large-scale information integration should be altered compared to the fully conscious state. This should also apply in systems with reduced consciousness, such as the cerebellum, and maybe organisms lower on the evolutionary tree.

We look forward to future work that confirms or falsifies our hypothesis.

References

- Alkire, M. T., McReynolds, J. R., Hahn, E. L. and Trivedi, A. N. (2007). Thalamic microinjection of nicotine reverses sevoflurane-induced loss of righting reflex in the rat. *Anesthesiology*, 107, 264–272.
- Birbaumer, N., Elbert, T., Canavan, A. G. and Rockstroh, B. (1990). Slow potentials of the cerebral cortex and behavior. *Physiological Reviews*, 70, 1–41.
- Braitenberg, V. and Schuz, A. (1998). *Cortex: Statistics and Geometry of Neuronal Connectivity*. Germany: Springer.
- Braun, A. R., Balkin, T. J., Wesenten, N. J., Carson, R. E., Varga, M., Baldwin, P., Selbie, S., Belenky, G. and Herscovitch, P. (1997). Regional cerebral blood flow throughout the sleep-wake cycle. An H₂(15)O PET study. *Brain*, 120, 1173–1197.

- Bullock, T. H. and Basar, E. (1988). Comparison of ongoing compound field potentials in the brains of invertebrates and vertebrates. *Brain Research*, 472, 57–75.
- Deecke, L. and Lang, W. (1988). P300 as the resolution of negative cortical DC shifts. *Behavioural Brain Research*, 11, 379–381.
- DeFelipe, J., Alonso-Nanclares, L. and Arellano, J. I. (2002). Microstructure of the neocortex: comparative aspects. *Journal of Neurocytology*, 31, 299–316.
- Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J. and Sergent, C. (2006). Conscious, pre-conscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences*, 10, 204–211.
- Douglas, R. J. and Martin, K. A. (2004). Neuronal circuits of the neocortex. *Annual Reviews in Neuroscience*, 27, 419–451.
- Fitzgerald, R. D., Lamm, C., Oczenski, W., Stimpfl, T., Vycudilik, W. and Bauer, H. (2001). Direct current auditory evoked potentials during wakefulness, anesthesia, and emergence from anesthesia. *Anesthesia and Analgesia*, 92, 154–160.
- Goldring, S. (1974). DC Shifts Released by Direct and Afferent Stimulation. In A. Remond (Ed.), *Handbook of Electroencephalography and Clinical Neurophysiology* (pp. 12–24). Amsterdam: Elsevier.
- Gray, H. M., Gray, K. and Wegner, D. M. (2007). Dimensions of mind perception. *Science*, 315, 619.
- Haggard, P. (2008). Human volition: towards a neuroscience of will. *Nature Reviews Neuroscience*, 9, 934–946.
- Hassler, R. (1978). Interaction of reticular activating system for vigilance and the thalamo-pallidal systems for directing awareness and attention under striatal control. In B. A. Rogeul-Buser (Ed.), *Cerebral Correlates of Conscious Experience* (pp. 111–129). Amsterdam; New York: Elsevier.
- He, B. J. and Raichle, M. E. (2009). The fMRI signal, slow cortical potential and consciousness. *Trends in Cognitive Sciences*, 13, 302–309.
- He, B. J., Snyder, A. Z., Zempel, J. M., Smyth, M. D. and Raichle, M. E. (2008). Electrophysiological correlates of the brain's intrinsic large-scale functional architecture. *Proceedings of the National Academy of Sciences of the USA*, 105, 16039–16044.
- Jones, E. G. (1998). Viewpoint: the core and matrix of thalamic organization. *Neuroscience*, 85, 331–345.
- Jones, M. O., Martin, C., Boorman, L., Harris, S., Bartlett, K., Kennerley, A., Zheng, Y., Berwick, J. and Mayhew, J. E. W. (2007). Interactions between evoked and spontaneous hemodynamic fluctuations in rodent brain. In 37th Annual Meeting of Society for Neuroscience, San Diego (San Diego).
- Khader, P., Schicke, T., Roder, B. and Rosler, F. (2008). On the relationship between slow cortical potentials and BOLD signal changes in humans. *International Journal of Psychophysiology*, 67, 252–261.
- Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I. and Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*, 320, 110–113.
- Laureys, S. (2005). The neural correlate of (un)awareness: lessons from the vegetative state. *Trends in Cognitive Sciences*, 9, 556–559.
- Leopold, D. A., Aura, C. and Maier, A. V. (2008). Laminar analysis of local field and current source density during physical and perceptual events in monkey V1. In 38th Annual Meeting of Society for Neuroscience, Washington D.C. (Washington, DC).

- Libet, B. (2004). *Mind Time: The Temporal Factor in Consciousness*. Cambridge, MA: Harvard University Press.
- Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, 453, 869–878.
- Mitzdorf, U. (1985). Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. *Physiology Reviews*, 65, 37–100.
- Nagai, Y., Critchley, H. D., Featherstone, E., Fenwick, P. B., Trimble, M. R. and Dolan, R. J. (2004). Brain activity relating to the contingent negative variation: an fMRI investigation. *Neuroimage*, 21, 1232–1241.
- Pins, D. and Ffytche, D. (2003). The neural correlates of conscious vision. *Cerebral Cortex*, 13, 461–474.
- Roland, P. E., Hanazawa, A., Undeman, C., Eriksson, D., Tompa, T., Nakamura, H., Valentiniene, S. and Ahmed, B. (2006). Cortical feedback depolarization waves: a mechanism of top-down influence on early visual areas. *Proceedings of the National Academy of Sciences of the USA*, 103, 12586–12591.
- Rosler, F., Heil, M. and Roder, B. (1997). Slow negative brain potentials as reflections of specific modular resources of cognition. *Biological Psychology*, 45, 109–141.
- Schiff, N. D., Giacino, J. T., Kalmar, K., Victor, J. D., Baker, K., Gerber, M., Fritz, B., Eisenberg, B., Biondi, T., O'Connor, J., Kobylarz, E. J., Farris, S., Machado, A., McCagg, C., Plum, F., Fins, J. J. and Rezai, A. R. (2007). Behavioural improvements with thalamic stimulation after severe traumatic brain injury. *Nature*, 448, 600–603.
- Sergent, C., Baillet, S. and Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience*, 8, 1391–1400.
- Soon, C. S., Brass, M., Heinze, H. J. and Haynes, J. D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, 11, 543–545.
- Tononi, G. and Koch, C. (2008). The neural correlates of consciousness: an update. *Annals of the New York Academy of Sciences*, 1124, 239–261.
- Vincent, J. L., Patel, G. H., Fox, M. D., Snyder, A. Z., Baker, J. T., Van Essen, D. C., Zempel, J. M., Snyder, L. H., Corbetta, M. and Raichle, M. E. (2007). Intrinsic functional architecture in the anaesthetized monkey brain. *Nature*, 447, 83–86.
- von Stein, A. and Sarnthein, J. (2000). Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. *International Journal of Psychophysiology*, 38, 301–313.

Glossary

In this article, we use “consciousness” or “conscious awareness” synonymously as “subjective awareness”. We use “conscious experience” to refer to the experience of subjective awareness. Lastly, “conscious state” refers to the physiological states under which conscious awareness is present.

Box 1. Brain networks and consciousness

Not all brain networks contribute to consciousness equally. We speculate that the anterior cingulate and anterior insular cortices, as well as the default network (a set of brain regions routinely decreasing their activity during goal-directed tasks (Vincent et al., 2007)), might be more pivotal than the sensory/motor networks and maybe even the dorsal attention network (including the dorsal visual stream and frontal eye field) in the emergence of consciousness. This conjecture mainly comes from a thought experiment comparing the largely unconscious state – slow-wave sleep (SWS), with the conscious states including wakefulness and rapid-eye-movement (REM) sleep. Whereas the sensory/motor regions and the dorsal attention network are as active in SWS as in wakefulness; the anterior cingulate, anterior insular and the midline regions of the default network are deactivated in SWS and reactivated in both REM sleep and wakefulness (Braun et al., 1997). To the best of our knowledge, this conjecture is also consistent with existing data from persistent vegetative patients, blindsight patients, and from manipulations of momentary conscious perception.

Distinct characteristics of conscious experience are met by large-scale neuronal synchronization

Lucia Melloni and Wolf Singer

Conscious processes have a number of distinct properties that need to be accounted for by neuronal mechanisms supporting conscious experience. It is proposed that synchronization of distributed neuronal activity patterns meets most of these requirements. A major problem for the identification of neuronal correlates of consciousness is the distinction between the processes that lead to conscious experience and those that follow once contents have become conscious. Criteria for this distinction are discussed in the context of published evidence.

Everybody knows what it is like to be conscious. However, even though all of us seem to have at least a rough idea about what consciousness is and where it is located (the brain), a universally accepted definition is still lacking. In the meantime, we hope that an exhaustive definition and a better characterization of consciousness will emerge as theories are complemented with scientific discoveries. The apparent lack of consensus regarding the definition of consciousness has not precluded advancement in the field of the scientific study of consciousness. Since the seminal paper of Crick and Koch (1990), several theoretical proposals (Dehaene et al. 2006; Lamme, 2006; Singer, 2000) as well as important empirical contributions regarding the quest for the neuronal correlates of consciousness have been made. However, some important questions remain unsolved. In the following we will put forward the idea that the search for a neuronal *correlate* of consciousness entails in fact the search for a neuronal *mechanism* that can account for several aspects of consciousness. In what follows we will first briefly discuss several characteristics of consciousness. Then we will discuss the proposal that neuronal synchrony between distant neuronal assemblies might be a key mechanism for perceptual awareness. Empirical evidence supporting neuronal synchrony as a key correlate of consciousness will be provided. Finally, some pressing questions in the field of consciousness research will be discussed.

Some basic definitions

In what follows we will understand awareness as the representation of internal or sensory changes that are accessible to introspection or direct report, and operationally we will consider it as being present whenever a subject can report the presence or absence of a stimulus (detection) or its identity (discrimination). Failure to detect or identify will be taken as sign of unawareness. Even though this is not an exhaustive definition, it allows for an objective evaluation of subjective data, a scientific program called “heterophenomenology” (Dennett, 1992). Given the difficulties in establishing causal relations, consciousness research has focused on the search for neuronal correlates that are defined as the minimal set of neuronal events which are necessary and/or sufficient for perceptual awareness (Chalmers, 2000). This definition, however, confounds the search for a mechanism with the identification of mere correlates. For example, an event-related potential or BOLD signal in a certain area cannot explain how consciousness arises because they are only measures. What is needed instead is a model that causally relates certain neuronal mechanism to consciousness, and to induce predicted changes of consciousness by interfering with these mechanisms. In the following we define several requirements that need to be met by a mechanism subserving consciousness, and then discuss whether neuronal synchronization fulfills these premises.

Properties of consciousness constraining neuronal implementations

One central property of perceptual awareness is its unified character. While our experience is unified, the architecture of the brain is distributed, consisting of multiple, specialized but highly interconnected processing areas. This raises the question, how the unity of conscious experience can arise from the distributed processing of the brain?

Furthermore, the contents of our experience constantly change over time, but at any given moment, they are discrete and distinct. Thus, the neuronal implementation of consciousness has to meet these requirements for a seamless flow of ever changing contents that are experienced as coherent in space and time.

At any moment in time, only a fraction of the ongoing processes in the brain gains access to consciousness. Thus, inherent to being conscious of something is a selection process that gates access to consciousness. This raises the question what distinguishes processes that enter consciousness from those that do not. As most results of neuronal processing have in principle access to consciousness, – with the exception of certain interoceptive signals – the distinguishing signature cannot be content specific but should be assignable in a dynamic way to all different

activity patterns that can give rise to conscious experience. Evidence indicates that processing results which are not represented in awareness, can still have an impact on behavior. These unconscious processes are not restricted to low level analyses of sensory signals but can include deep semantic decoding, the programming of motor responses and they can even exert inhibitory control (Dehaene et al., 1998; van Gaal et al., 2008). Thus, theories of the neuronal correlates of consciousness have to address the question how signals are selected for access to awareness and which signatures distinguish conscious from unconscious information processing.

Neuronal synchrony as a key correlate of perceptual awareness

We shall argue that neuronal synchrony possesses most of the features that we have defined above as constitutive for consciousness or for processes leading to conscious awareness.

In order to create a unified conscious experience of multimodal contents, the modular architecture of the brain has to be overcome. One solution would be convergence of all information in a singular center. This option is considered theoretically implausible (Dennett, 1992), and so far no such center has been identified. Furthermore, behavioral and brain imaging studies have shown that unconscious processing engages very much the same cerebral areas as conscious processing, including frontal and prefrontal cortex (Lau and Passingham, 2007; van Gaal et al., 2008). Thus, there is no compelling evidence for specific areas supporting conscious processing. An alternative to convergence is that the distributed processes are bound together dynamically by rendering the signals coherent in time (Hebb, 1949), e.g. by synchronization (Singer and Gray, 1995). In this framework, the presence or absence of a particular feature is signaled through the activity of feature-selective neurons, while the grouping of those elementary features is achieved via synchronization of the respective neurons into functionally coherent assemblies, forming a distributed representation of a perceptual object. Neuronal synchronization is a self-organizing process that allows rapid formation and dissolution of neuronal assemblies defined by coherence. Such dynamic binding is an economical and flexible strategy to cope with the representation of the virtually unlimited variety of feature constellation characterizing perceptual objects. Taking the unified nature of conscious experience and the diversity of possible contents into account, coherence (synchrony) offers itself as a mechanism allowing representation of ever changing constellations of content in a unifying format. In this way the assemblies representing the different contents appearing as unified in the continuous flow of consciously processed content could be bound together into a coherent whole (meta-assembly) that is subject to continuous change by

the cooption and exclusion of subassemblies. In this framework the rate limiting factor for the formation of a new meta-assembly corresponds to the time needed to establish stable coherence. In case coherence is expressed by synchrony, this would be the time needed to establish stable phase relations. Stable states would then be reached once the relative phase between local oscillations ceases to change (Tognoli and Kelso, 2009).

Synchronization is also ideally suited to contribute to the selection of contents for access to consciousness. Synchronization enhances the saliency of signals and thereby facilitates their propagation in sparsely connected networks such as the cerebral cortex (Abeles, 1991; Fries, 2009; Jensen et al., 2007). Gamma band synchronization, in particular, assures coincidence among distributed inputs with millisecond precision. Furthermore, when neuronal responses engage in synchronized oscillations, frequency and phase adjustments can be exploited for the selective routing of activity and the dynamic gating of interactions between interconnected neurons. At the level of individual neurons oscillations are associated with periodic alternations of phases with high and low excitability, the latter resulting from the barrage of synchronized IPSPs that have both a shunting and a hyperpolarizing effect (Schroeder et al., 2008). Excitatory input that arrives at the depolarizing slope of an oscillation cycle generate large responses, while inputs arriving at the falling slope and trough are shunted and ineffective. Hence, neuronal oscillations define temporal windows for effective communication between neurons, providing a mechanism to selectively and flexibly bias the communication between neuronal groups (Fries, 2009). When two groups of neurons open their windows of susceptibility (their excitatory phases) at the same time, they are more likely to interact with each other, to increase their synchrony and as a consequence to also have enhanced influence on other groups of neurons. By adjusting oscillation frequency and phase, groups of neurons can either be linked together into tightly coupled assemblies or be segregated into functionally isolated subgroups. This mechanism can act both within and across areas, and can in principle account for selective and flexible routing of information within networks with fixed anatomical architectures. Taken together, the oscillatory patterning of activity and the option to adjust frequency and phase of the oscillations could serve three complementary functions: gain control, selective and flexible routing of information between neuronal groups, and formation of coherent representations. Furthermore, if depth of processing is determined by the extent of propagation of information in the brain, this can also account for the observation that conscious perception is associated with deeper processing than unconscious perception.

As previously mentioned, any neuronal correlate of consciousness should distinguish between conscious and unconscious forms of information processing.

Regarding neuronal synchrony, this distinction can be satisfied when considering its spatial scale: Following Varela et al.'s (2001) distinction between local and global scale integration, it can be postulated that processing carried out unconsciously (which is automatic and modular) relies on local integration in divergent-convergent feed-forward architectures, whereas large scale integration via reciprocal pathways subserves conscious processing (for a similar proposal see Dehaene et al., 2006). Anatomical and physiological studies suggest that neuronal, in particular cortical architectures share features of “small world networks” (reviewed in Bassett and Bullmore, 2006). These allow for the coexistence of both local, modular and global, distributed processes. Important properties of this architecture are minimization of path length between any pairs of nodes, optimization of the number of connections and the possibility of coexistence of multiple local processes and globally ordered states. Moreover, such networks can operate in critical states, allowing for fast reconfigurations of network dynamics (Bassett and Bullmore, 2006; Sporns and Zwi, 2004; Yu et al., 2008). Cortico-cortical connections can be subdivided in two major subgroups. Local, intracortical connections that run tangentially to the layers and link neurons that share similar response properties and are separated by only a few hundred micrometers, and long distance connections that often but not always run through the underlying white matter and link neurons in different cortical areas. The latter serve transmission of information to distinct cortical sites and can establish globally coordinated activation patterns (Varela et al., 2001).

We started this chapter describing four conditions that any theory of the neuronal correlates of consciousness should satisfy. We then discussed how neuronal synchronization could fulfill each of them, namely that neuronal synchronization could account for (i) the unity of experience as well as for (ii) its temporal dynamics, (iii) for the selection of signals entering conscious processing and for (iv) the distinction between conscious and unconscious processing related to the spatial scale of the synchronized activity. In what follows we will review evidence relating neuronal synchronization to consciousness.

Evidence relating long-range synchronization and consciousness

Masking offers an interesting possibility to study consciousness since the same physical stimuli can be either perceived or not perceived depending on the temporal and spatial sequence of stimuli that surround them. In our studies, we capitalized on this phenomenon and presented words that could be perceived in some trials and not in others (by adjusting the luminance of the mask) and simultaneously performed electroencephalographic (EEG) recordings (Melloni et

al., 2007). Several measures were analyzed: Time-resolved power changes of local signals, the precision of phase synchronization across recording sites over a wide frequency range, and event-related potentials (ERPs). A brief burst of long distance synchronization in the gamma frequency range between occipital, parietal and frontal sensors was the first event that distinguished seen from unseen words, while local synchronization was similar between conditions. Interestingly, after this transient period of synchronization, several other measures differed between seen and unseen words: We observed an increase in amplitude of the P300 ERP for visible words which most likely corresponds to the transfer of information to working memory. In addition, during the interval period in which visible words had to be maintained in memory, we observed increases in frontal theta oscillations. Theta oscillations have been related to maintenance of items in short-term memory (Jensen and Tesche, 2002).

To test whether the increase in long-distance synchronization relates to awareness or depth of processing, we further manipulated the depth of processing of invisible words. It has previously been shown that invisible words can be processed up to the semantic and motor level (Dehaene et al., 1998). In a subliminal semantic priming experiment we briefly presented words (invisible) that could either be semantically related or not related with a second visible word on which subjects had to carry out a semantic classification task. Invisible words were processed up to semantic levels as revealed by modulation of the reaction times depending on the congruency between invisible and visible words: congruent pairs exhibited shorter reaction times than incongruent ones. We observed increases in power in the gamma frequency range for unseen but processed words. For visible words we additionally observed increases in long-distance synchronization in the gamma frequency range (Melloni and Rodriguez, 2007). Thus, local processing of stimuli is reflected in increases in gamma power, whereas long-distance synchronization seems to be related to awareness of the stimuli. This suggests that conscious processing requires a particular dynamical state of the cortical network. The large-scale synchronization that we observed in our study could reflect the transfer of contents into awareness and/or their maintenance. We favor the first possibility given the transient nature of the effect and argue that the subsequent theta oscillations might support maintenance. It is conceivable that short periods of long-distance synchronization in the gamma band reflect the update of new contents, while the slower pace of theta oscillations might relate to sustained integration and maintenance of local results. The interplay between these two frequency bands might underlie the phenomenon of continuous but ever changing conscious experience (see below).

Recently, Gaillard et al. (2009) revisited the question of processing of visible and invisible words. In intracranial recordings in epileptic patients they observed

that invisible words elicited activity in multiple cortical areas which quickly vanished after 300 ms. In contrast, visible words elicited sustained voltage changes, increases in power in the gamma band, as well as long-distance synchronization in the beta band and long-range Granger causality. In contrast to our study, Gaillard et al. observed a rather late (300–500 ms) rise of long-distance synchronization. However, it is important to note that in the study of Gaillard et al., phase-synchrony was analyzed mostly over electrodes within a given cortical area or at most between hemispheres, and thus this analysis raises an important coverage issue. Earlier synchronization events could have passed undetected because of lack of electrode coverage. Since the electrode placement in such studies is based on medical and not experimental considerations, analyses are restricted to the available electrodes and locations. Despite of these restrictions, this study provides one of the most compelling pieces of evidence for a relation between long-distance synchronization and consciousness.

Another commonly used paradigm in studies of consciousness is binocular rivalry. When two images that cannot be fused are presented to each of the two eyes, perception fluctuates between the two images. Thus, despite constant stimulation conditions, perception alternates. This suggests a selection mechanism that gates access to consciousness. Several studies using different stimulus materials as well as recording methods (single cell analysis and LFP recordings in animals and EEG and MEG registration in human subjects) have shown increased synchronization and phase locking of oscillatory responses to the stimulus that was consciously perceived and controlled behaviour (Cosmelli et al., 2004; Fries et al., 1997; Srinivasan et al., 1999). Cosmelli et al. (2004) extended the findings obtained in human subjects by performing source reconstruction and analyzing phase-synchrony in source space. These authors observed that perceptual dominance was accompanied by coactivation of occipital and frontal regions, including anterior cingulate and medial frontal areas. Recently, Doesburg et al. (2009) provided evidence for a relation between perceptual switches in binocular rivalry and theta and gamma band synchronization. Perceptual switches were related to increments in long-distance synchronization in the gamma band between several cortical areas (frontal and parietal) that repeated at the rate of theta oscillations. The authors suggested that transient gamma-band synchronization supports discrete moments of perceptual experience while theta oscillations structure their succession in time, pacing the formation and dissolution of distributed neuronal assemblies. Thus, long-range gamma synchronization locked to ongoing theta oscillations could serve to structure the flow of conscious experience allowing for changes in content every few hundred millisecond. Further research is required to clarify the exact relation between the two frequency bands and their respective role on the generation of percepts and the pacing of changes in perception.

Another paradigm in consciousness research exploits the attentional blink phenomenon. When two stimuli are presented at short intervals among a set of distractors, subjects usually detect the first (S1) but miss the second (S2) when it is presented 200–500 ms after S1. Increases in long-range neuronal synchrony in the beta and gamma frequency ranges have been observed when the S2 is successfully detected (Gross et al., 2004; Nakatani et al., 2005). Furthermore, Gross et al. (2004) observed that successful detection of both S1 and S2 was related to increased long-distance synchronization in the beta range to both stimuli, and this enhanced synchrony was accompanied by higher desynchronization in the inter-stimulus-interval. Thus, desynchronization might have facilitated the segregation of the two targets, allowing for identification of the second stimulus (also see Rodriguez et al., 1999). Source analysis revealed, as in the case of binocular rivalry, dynamical coordination between frontal, parietal, and temporal regions for detected targets (Gross et al., 2004).

In summary, studies of masking, binocular rivalry, and the attentional blink support the involvement of long-range synchronization in conscious perception. Recent investigations have suggested further that a nesting of different frequencies, in particular of theta and gamma oscillations, could play a role in pacing the flow of consciousness. Furthermore, the study of Gross et al. (2004), suggests that desynchronization could serve to segregate representations when stimuli follow at short intervals. These results are encouraging and should motivate further search for relations between oscillatory activity in different frequency bands and consciousness, whereby attention should be focused not only on the formation of dynamically configured networks but also on their dissolution.

Pressing questions in the field of consciousness research

One influential view on the function of consciousness posits that consciousness allows for exchange of contents between several independent processing modules, which in turn makes this information available to a variety of processes including perceptual categorization, unification in short and long-term memory, evaluation, etc. (Baars, 1997). If consciousness is a prerequisite of these cognitive processes, how can the requirements of and the consequences of consciousness be distinguished? If conscious and unconscious processing did not only differ with respect to their respective qualitative properties, but also with respect to their consequences, such a distinction might indeed be impossible: If consciousness, as proposed, enhances the depth of processing then a method that contrasts differences in perceptual states of physically identical stimuli will reveal not only activity related to awareness *per se* but also activity related to its consequences,

i.e. deeper processing, episodic memory formation, etc. (for a similar argument see Lau, 2008). This might explain why different research groups find signatures of consciousness at different moments in time. Neurons in the medial temporal lobe (MTL) show an all or none response profile depending on whether stimuli are processed consciously. When briefly presented images are made unperceivable through backward masking, neurons in MTL do not respond. In contrast, when the same images are recognized, MTL neurons exhibit clear responses with a latency of about 300 ms (Quiroga et al., 2008). Following the logic of the contrastive approach, this result can be taken as a correlate of consciousness. The question is: is this a correlate of consciousness or a correlate of the consequences of consciousness? Given the intimate relation between MTL and hippocampus and hence with networks responsible for the management of episodic memory it is also conceivable that the rather late MTL responses have to do with memory formation following conscious perception rather than with the access to consciousness per se. This view is supported by the fact that complete resection of the hippocampus and adjacent temporal cortex does not lead to deficits in conscious perception (Postle, 2009). However, this does also not imply that MTL does not contribute to conscious perception. It is important to distinguish between neuronal signatures which correlate with the content of our moment-to-moment conscious awareness, and those that contribute to the feeling of a stream of consciousness (continuous present). Thus, MTL activity might not contribute to the former, but be essential to the latter. In order to experience a continuous present, memories from events that have just occurred should be linked to those occurring at the present time. Thus, the delayed MTL activity could serve to assure the continuity of the flow of consciousness.

How else could one solve the conundrum to distinguish between cause and effect? There is no panacea. Obviously, the most straightforward approach is to obtain comprehensive data on the sequence of effects distinguishing conscious from unconscious processing, assuming that causes precede effects. This could be complemented with attempts to interfere with these sequential processing using TMS. For instance, if blocking the late activity in MTL does not abolish conscious perception but its continuity we would be one step further. Following this strategy those processes could be discarded one by one that figure as consequences of consciousness leaving us with those more closely related to consciousness per se. Secondly, one could evaluate introspective reports as a direct measure of the quality of the experience. In this case one would not contrast conscious with unconscious processing, but arrive at a parametric estimate of “consciousness”. For instance, in visual experiments we could ask subjects about how clear their content of awareness is, and then correlate increases in subjective clarity with measures of neuronal activity.

Conclusions

It is evident that the study of consciousness has greatly profited from the search for neuronal correlates. However, simply showing that the brain makes a difference between conscious and unconscious processing is not satisfactory. We propose here to go one step further and to develop mechanistic explanations that establish plausible, and hopefully at later stages, even causal relations between brain processes and consciousness. Oscillatory synchrony is one candidate mechanism, and it has the advantage that it can be measured relatively directly in humans which are able to give detailed descriptions about their conscious experience. However, oscillations and synchrony seem to be mechanisms that are as intimately and inseparably related to neuronal processing in general just as the modulation of neuronal discharge rates. Thus, without further specification these phenomena cannot stand up as a neuronal correlates of consciousness apart from the triviality that consciousness does not exist without them. We and others (Varela et al., 2001) propose that the spatial scale and perhaps also the precision and stability of neuronal synchrony might be taken as more specific indicators of whether the communication of information in the brain is accompanied by conscious experience or not. In this framework, conscious experience arises only if information that is widely distributed within or across subsystems is not only processed and passed on to executive structures but bound together into a coherent, all-encompassing, non-local but distributed meta-representation. This interpretation is compatible with views considering consciousness as the result of the dynamic interplay of brain subsystems; but it poses the challenging question related to the “hard problem of consciousness” research: Is there something on top of this distributed meta-assembly that makes us experience? From a truly dynamicist point of view, the answer is “probably not”. However, this should not discourage us from trying to separate the contribution of each subsystem and then go on to characterize how they interact. This would be a real alternative to the ever lurking homunculus.

Acknowledgements

This work was supported by the Max Planck Society. We are indebted to Caspar M. Schwiedrzik for insightful comments on this manuscript.

References

- Abeles, M. (1991). *Corticotronics: Neural circuits of the cerebral cortex*. Cambridge, MA: Cambridge University Press.
- Baars, B. J. (1997). In *the theatre of consciousness: The workspace of the mind*. New York, NY: Oxford University Press.
- Bassett, D. S. and Bullmore, E. (2006). Small-world brain networks. *Neuroscientist*, 12, 512–523.
- Chalmers, D. J. (2000). What is a neural correlate of consciousness? In T. Metzinger (Ed.), *Neural correlates of consciousness: Empirical and conceptual questions* (pp. 17–40). Cambridge, MA: MIT Press.
- Cosmelli, D., David, O., Lachaux, J. P., Martinerie, J., Garnero, L., Renault, B. and Varela, F. (2004). Waves of consciousness: Ongoing cortical patterns during binocular rivalry. *Neuroimage*, 23, 128–140.
- Crick, F. and Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in Neuroscience*, 2, 263–275.
- Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J. and Sergent, C. (2006). Conscious, pre-conscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences*, 10, 204–211.
- Dehaene, S., Naccache, L., Le Clec, H. G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., van de Moortele, P. F. and Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature*, 395, 597–600.
- Dennett, D. C. (1992). *Consciousness explained*. London: Penguin.
- Doesburg, S. M., Green, J. J., McDonald, J. J. and Ward, L. M. (2009). Rhythms of consciousness: Binocular rivalry reveals large-scale oscillatory network dynamics mediating visual perception. *PLoS One*, 4, e6142.
- Fries, P. (2009). Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annual Reviews in Neuroscience*, 32, 209–224.
- Fries, P., Roelfsema, P. R., Engel, A. K., König, P. and Singer, W. (1997). Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry. *Proceedings of the National Academy of Sciences of the USA*, 94, 12699–12704.
- Gaillard, R., Dehaene, S., Adam, C., Clemenceau, S., Hasboun, D., Baulac, M., Cohen, L. and Naccache, L. (2009). Converging intracranial markers of conscious access. *PLoS Biol*, 7, e61.
- Gross, J., Schmitz, F., Schnitzler, I., Kessler, K., Shapiro, K., Hommel, B. and Schnitzler, A. (2004). Modulation of long-range neural synchrony reflects temporal limitations of visual attention in humans. *Proceedings of the National Academy of Sciences of the USA*, 101, 13050–13055.
- Hebb, D. O. (1949). *The organization of behavior*. New York, NY: Wiley.
- Jensen, O., Kaiser, J. and Lachaux, J. P. (2007). Human gamma-frequency oscillations associated with attention and memory. *Trends in Neurosciences*, 30, 317–324.
- Jensen, O. and Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *European Journal of Neuroscience*, 15, 1395–1399.
- Lamme, V. A. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, 10, 494–501.

- Lau, H. C. (2008). Are we studying consciousness yet? In L. Weiskrantz and M. Davies (Eds.), *Frontiers of Consciousness. The Chichele Lectures* (pp. 245–258). Oxford: Oxford University Press.
- Lau, H. C. and Passingham, R. E. (2007). Unconscious activation of the cognitive control system in the human prefrontal cortex. *Journal of Neuroscience*, 27, 5805–5811.
- Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W. and Rodriguez, E. (2007). Synchronization of neural activity across cortical areas correlates with conscious perception. *Journal of Neuroscience*, 27, 2858–2865.
- Melloni, L. and Rodriguez, E. (2007). Non-perceived stimuli elicit local but not large-scale neural synchrony. *Perception*, 36 (ECP Abstract Supplement).
- Nakatani, C., Ito, J., Nikolaev, A. R., Gong, P. and van Leeuwen, C. (2005). Phase synchronization analysis of EEG during attentional blink. *Journal of Cognitive Neuroscience*, 17, 1969–1979.
- Postle, B. R. (2009). The hippocampus, memory, and consciousness. In S. Laureys and G. Tononi (Eds.), *The Neurology of consciousness: Cognitive Neuroscience and Neuropathology* (pp. 326–338). London: Academic Press.
- Quiroga, R. Q., Mukamel, R., Isham, E. A., Malach, R. and Fried, I. (2008). Human single-neuron responses at the threshold of conscious recognition. *Proceedings of the National Academy of Sciences of the USA*, 105, 3599–3604.
- Rodriguez, E., George, N., Lachaux, J. P., Martinerie, J., Renault, B. and Varela, F. J. (1999). Perception's shadow: Long-distance synchronization of human brain activity. *Nature*, 397, 430–433.
- Schroeder, C. E., Lakatos, P., Kajikawa, Y., Partan, S. and Puce, A. (2008). Neuronal oscillations and visual amplification of speech. *Trends in Cognitive Sciences*, 12, 106–113.
- Singer, W. (2000). Phenomenal awareness and consciousness from a neurobiological perspective. In T. Metzinger (Ed.), *Neural correlates of consciousness: empirical and conceptual questions* (pp. 121–137). Cambridge, MA: MIT Press.
- Singer, W. and Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Reviews in Neuroscience*, 18, 555–586.
- Sporns, O. and Zwi, J. D. (2004). The small world of the cerebral cortex. *Neuroinformatics*, 2, 145–162.
- Srinivasan, R., Russell, D. P., Edelman, G. M. and Tononi, G. (1999). Increased synchronization of neuromagnetic responses during conscious perception. *Journal of Neuroscience*, 19, 5435–5448.
- Tognoli, E. and Kelso, J. A. (2009). Brain coordination dynamics: true and false faces of phase synchrony and metastability. *Progress in Neurobiology*, 87, 31–40.
- van Gaal, S., Ridderinkhof, K. R., Fahrenfort, J. J., Scholte, H. S. and Lamme, V. A. (2008). Frontal cortex mediates unconsciously triggered inhibitory control. *Journal of Neuroscience*, 28, 8053–8062.
- Varela, F., Lachaux, J. P., Rodriguez, E. and Martinerie, J. (2001). The brainweb: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, 2, 229–239.
- Yu, S., Huang, D., Singer, W. and Nikolic, D. (2008). A small world of neuronal synchrony. *Cerebral Cortex*, 18, 2891–2901.

Gamma oscillations and the cellular components of consciousness?

Fiona E. N. LeBeau

There is evidence to suggest that synchronised neuronal activity, particularly that in the gamma frequency range ($\sim 30\text{--}100\text{Hz}$) might contribute to conscious perception. In this chapter I shall discuss what we know about the mechanisms that underlie the generation of gamma frequency activity and the changes known to occur in diseases such as schizophrenia. Until recently it was difficult to test the role of individual neuronal subtypes in the generation of network oscillations. However, using the newly developed optogenetic techniques (in which cells are genetically modulated to make them responsive to specific wavelengths of light) it should now be possible to directly test the contribution of different neuronal classes to the generation of network oscillations and perhaps, therefore, consciousness.

Renee Magritte painted “Dreaming” in 1935 (Figure 1B). In the upper half of the picture a man lies asleep in a small chamber while below the random objects that appear in his dreams are illustrated below him. Lining the walls of the sleeping man’s chamber are a series of irregular lines that are remarkably similar to patterns of activity that emerge from the brain recorded by the electroencephalogram (EEG; Figure 1A).

The first recordings of human brain activity were made by Hans Berger, (a few years before Magritte painted “Dreaming”), when he recorded a 10 Hz rhythm called the alpha rhythm (Berger, 1929). I doubt Magritte knew about the brain’s electrical activity when he painted this picture but since 1925 it has been shown that the brain generates rhythmic activity over a wide range of frequencies from very low $<1\text{ Hz}$, to very fast 100s of Hz, and that different patterns of activity occur during different behavioural states – including dreaming. Different EEG rhythms, or so-called neuronal oscillations, occur during different states of consciousness with large slow delta ($\sim 1\text{--}4\text{ Hz}$) waves recorded during deep sleep while faster oscillations such as a beta ($\sim 15\text{--}30\text{ Hz}$) and gamma ($\sim 30\text{--}100\text{ Hz}$) are associated with the awake conscious state.

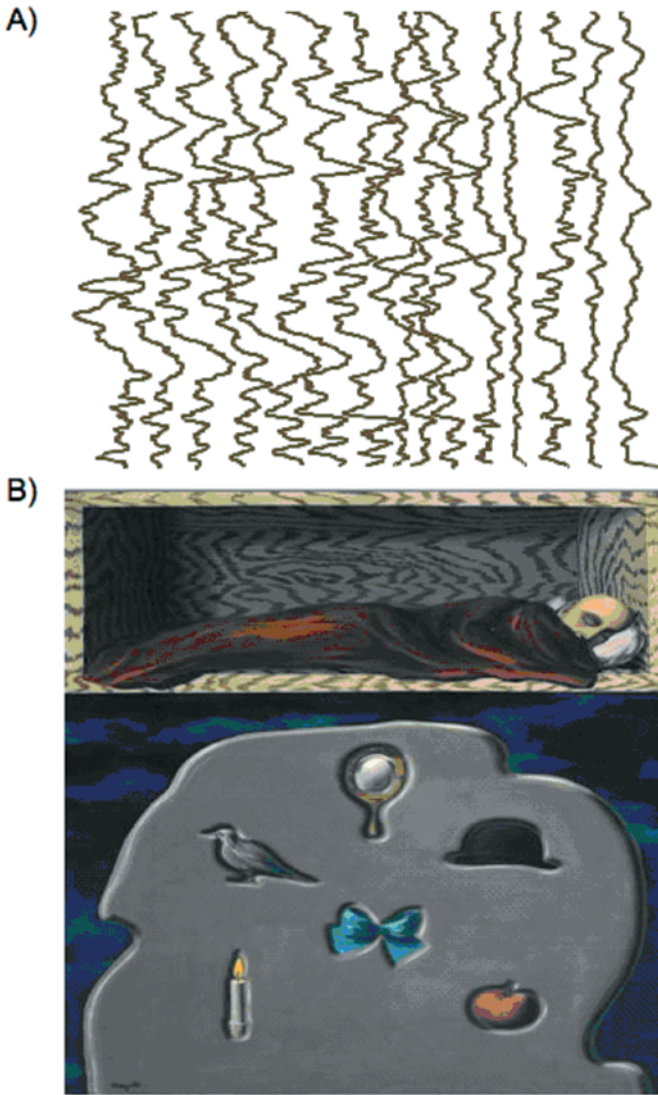


Figure 1. (A) Show sample EEG recordings from 15 electrodes from human subject during slow wave sleep. (B) Renee Margrite’s “Dreaming” 1935.

In particular, synchronous gamma frequency oscillations have been proposed to be an important neural correlate of consciousness (Engel et al., 1999; Fries et al., 1997; 2001; Dehaene et al., 2006). See also Melloni and Singer’s chapter for a detailed discussion of neuronal synchronisation and conscious experience. In

this chapter I will discuss the evidence for the role of gamma frequency activity in conscious perception and outline how we can, with new optogenetic technology, potentially identify the cellular components that are essential for consciousness.

Mechanisms underlying the generation of gamma frequency oscillations

Synchronous neuronal activity in the gamma frequency band has been recorded in a number of brain regions and has been linked to a range of cognitive functions including sensory processing (Engel et al., 1999), attention (Fries et al., 2001), and memory functions (Sederberg et al., 2007). Oscillatory activity can now be observed in EEG and magnetoencephalogram (MEG) recordings in humans and animals, and electrophysiological recordings can also be made in animals *in vivo*. An important advance in understanding the cellular mechanisms underlying the generation of gamma frequency activity came from the observation that this brain rhythm can also be generated in isolated brain slices *in vitro* (Whittington et al., 1995; Fisahn et al., 1998).

Cortical interneurons using the inhibitory neurotransmitter gamma-aminobutyric acid (GABA) have been shown, in both *in vitro* (Whittington et al., 1995; Fisahn et al., 1998) and *in vivo* (Klausberger et al., 2003) studies, to be essential for the generation of gamma frequency oscillations. While only constituting approximately 10–20% of neurons in the cortex detailed morphological and electrophysiological studies have revealed a vast complexity and diversity of cortical inhibitory interneurons. In the hippocampus, for example, ~20 different cell types have now been reported and it is clear that different classes of interneuron play distinct roles in the generation of different types of oscillatory activity (Somogyi and Klausberger, 2005; Grillner et al., 2005). No single criteria can be used for the classification of interneurons, but information about cell morphology, physiological response properties and molecular markers are combined to identify key subtypes (Freund and Buzsaki, 1996). Important neurochemical markers of interneuron subtype include the expression of distinct calcium binding proteins, such as parvalbumin, calbindin or calretinin, and the expression of neuropeptides such as somatostatin and cholecystokinin (Freund and Buzsaki, 1996).

Parvalbumin-containing interneurons in particular have been shown to be critical for the generation of gamma frequency activity. Parvalbumin-containing basket cells and chandelier cells provide powerful inhibitory input to the soma, and axon initial segment respectively, that serves to time the firing of pyramidal cells. In addition, the wide diversity of synaptic contacts allows these interneurons to co-ordinate activity in large populations of pyramidal cells and so synchronise

activity. In addition to chemical neurotransmission, electrical signalling via gap junctions between both interneurons and pyramidal cells, is also important for the generation of fast cortical oscillations (for review see LeBeau et al., 2003; Traub et al., 2004). These, and many other studies over the past 15 years, have gone a long way to identifying the key cellular components required for the generation of a network rhythm that may be important for consciousness.

Role of gamma frequency oscillations

Singer and colleagues (for review see Singer and Gray, 1995) first identified the role of synchronous neuronal gamma frequency oscillations in perception in the visual system. Since then gamma frequency oscillations have been seen in all sensory, motor, and frontal cortical areas as well as many other brain regions such as the hippocampus and amygdala. Gamma frequency oscillations were initially proposed to be important for “binding”, the linking together of activity in different neuronal populations that encode different components of a sensory input (i.e. colour or shape) into a single percept (Engel et al., 1999). Transient synchronisation of widely distributed neural networks was proposed to be necessary for conscious perception (Engel et al., 1999).

Recent studies have provided further evidence for a link between gamma frequency oscillations and conscious processing (Fries et al., 1997; 2001; Melloni et al., 2007). One study compared EEG activity in human subjects when presented with a series of word stimuli, some of which were consciously perceived, others were masked words that were processed, but remained non-conscious (Melloni et al., 2007; see also Melloni and Singer’s chapter for more detailed discussion). An increase in the phase locking, or synchrony, of gamma band activity across multiple regions was observed after the presentation of the word stimulus *only* when the word was consciously perceived (Melloni et al., 2007). The authors proposed that the transient global synchrony observed ultimately permitted access to consciousness. Dehaene and colleagues (Dehaene et al., 2006) had previously suggested non-conscious processing was related to local changes in co-ordinated activity, but conscious perception required global synchrony across widely distributed brain regions.

However, gamma frequency activity cannot simply be equated with conscious sensory perception. Gamma frequency activity can be recorded in the EEG in a range of different states of consciousness including during the REM phase of sleep (Tonnoni and Koch, 2008), during hallucinations in patients with schizophrenia (Baldeweg et al., 1998) and in patients with tinnitus (Ashton et al., 2007). We do not yet know if these represent the *same* gamma frequency activity, in that the

identical neuronal components are involved in each of these *types* of gamma frequency activity. Future work (discussed below) may, however, enable us to further sub-divide gamma frequency activity based on the cellular components/brain regions that underlie its generation.

Changes in gamma frequency oscillations in disease

If gamma frequency oscillations are essential for normal conscious perception then we should expect to see changes in gamma frequency oscillations in conditions of altered conscious perception. Furthermore, understanding how network synchrony has changed in a range of neuropathological and neuropsychological conditions may help elucidate the cellular changes that occur in these conditions. Abnormalities in gamma frequency activity have now been reported in Alzheimer's disease, schizophrenia and, autism (see Uhlhaas and Singer, 2006 for review).

Of these conditions schizophrenia has been the most extensively studied. Patients with schizophrenia exhibit a range of perceptual abnormalities, most notably in visual perception and cognitive deficits (see Lewis and Gonzales-Burgos, 2008; Roopun et al., 2008 for reviews). In addition, several studies have reported impaired gamma frequency activity in patients with schizophrenia (Spencer et al., 2003; Uhlhaas and Singer, 2006). A reduced visual gamma band response (Spencer et al., 2003) and a reduction in long-range synchrony in the gamma band have both been reported in schizophrenic patients (Uhlhaas and Singer, 2006). These deficits in oscillatory activity could reflect either a change in the local neuronal circuit and/or a change in long-range connectivity.

One key common pathological feature associated with conditions of altered consciousness such as schizophrenia is a reduction in GABA inhibitory neurotransmission (see Lewis and Gonzales-Burgos, 2008 for review). In particular two key pathological changes have been well documented. Firstly, a decrease in glutamic acid decarboxylase (GAD-67) the enzyme responsible for synthesising GABA protein has been observed in the cortex in human post-mortem studies from patients with schizophrenia. Secondly, there is a loss of parvalbumin labelling in interneurons but no cell loss; although other interneuron classes may also be affected (see Lewis and Gonzales-Burgos, 2008 for review).

Changes in the inhibitory network might, therefore, account for the changes in gamma frequency activity observed in patients with schizophrenia (Spencer et al., 2003; Uhlhaas and Singer, 2006). Working memory (the ability to transiently maintain and manipulate a limited amount of information to guide thought and behaviour) deficits are a key feature of the cognitive deficits associated with

schizophrenia. It has been suggested that altered GABA neurotransmission in the prefrontal cortex may cause impaired gamma frequency and subsequently impaired working memory function in patients with schizophrenia (see Lewis and Gonzales-Burgos, 2008 for review).

Do changes in parvalbumin-expressing interneurons also account for the changes in gamma frequency activity observed in other conditions? More detailed studies are currently required to answer this question. However, although there are conflicting reports in the literature, reductions in parvalbumin labelling have now been reported in human post-mortem studies in patients with Alzheimer's disease (Brady et al., 1997) and dementia with Lewy bodies (Bernstein et al., 2010). Reduced GABAergic neurotransmission, including reduced GAD-67 expression, have also been reported in autism in which deficits in gamma frequency activity similar to those reported in schizophrenia occur (see Uhlhaas and Singer, 2006).

Optogenetics – A tool to modulate consciousness

As discussed above there is evidence to link changes in parvalbumin-expressing interneurons to changes in network oscillations and thus perhaps cognitive function and some form of conscious experience. However, to uniquely identify the role of individual cell types in consciousness new approaches are needed. While we have learnt a considerable amount about brain function and network oscillations from electrophysiological recording techniques, most notably extracellular recordings of spikes and field oscillations, the techniques are limited and it is difficult to identify the cells that have been recorded. Some approaches use analytical techniques to identify the broad class of cell (excitatory versus fast spiking interneuron) on the basis of the extracellular spike waveform. Others have combined extracellular recording with juxtacellular labelling of the cell with biocytin (Klausberger et al., 2003) but this is technically challenging and the yield of identified cells is low. A further limitation of conventional electrophysiological techniques is that electrical stimulation, which is used to activate neuronal networks, will excite all cells in the vicinity of the electrode indiscriminately. Key to understanding the cellular components required for consciousness will be our ability to manipulate individual neuronal elements and determine their effects on behaviour. In particular, we will need to be able to specifically activate individual classes of cells in a complex neuronal network, in specific targeted brain regions. In the past few years the advent of optogenetic recordings techniques has provided the means to do just that (Boyden et al., 2005; Han and Boyden, 2007; Cardin et al., 2009; Sohal et al., 2009).

Optogenetic techniques use naturally occurring microbial opsin molecules that can be activated or silenced by pulses of light of the appropriate wavelength. These molecules can be expressed in neurons in the brain using viral infection or transfection techniques. Two widely used opsin molecules activated by different wavelengths of light are *Natromonas pharaonis* halorhodopsin (eNpHR) – a light-driven chloride pump from Archaeobacterium and channelrhodopsin 2 (ChR2) – a light-gated cation channel from algae (Boyden et al., 2005; Han and Boyden, 2007; Cardin et al., Sohal et al., 2009). Different opsin molecules can be specifically targeted to either pyramidal cells or interneurons and, when activated by a brief pulse of light of the appropriate wavelength, can cause either excitation or inhibition of neuronal responses.

Two recent studies have highlighted the role of parvalbumin-containing interneurons in the generation of gamma frequency activity (Cardin et al., 2009; Sohal et al., 2009). In one study Sohal et al., (2009) targeted the expression of the two different opsin molecules (eNpHR and ChR2) to parvalbumin-containing interneurons in the hippocampus. Reducing activity in the parvalbumin-containing interneurons *in vivo* resulted in a decrease in gamma frequency activity while activating these interneurons was sufficient to evoke gamma frequency activity (Sohal et al., 2009). In the second study Cardin et al., (2009) specifically labelled parvalbumin-containing interneurons and pyramidal cells in the mouse barrel cortex. They found that activation of the fast-spiking interneurons *in vivo* with light pulses at 20–80Hz amplified gamma frequency activity. These two studies thus confirmed the role of parvalbumin-containing interneurons in the generation of gamma band activity.

Future work in which individual cell types in a network can be switched on, or off, in an awake behaving condition using optogenetic techniques means it should be possible to test for the identify of the key cellular requirements for consciousness. For example, questions such as what cell types are needed, where are these cells located, how many cells need to be active, can all be addressed using these techniques (Han and Boyden, 2007). Using laser optical fibres coupled to electrodes allows one to activate or inhibit specific cell types in multiple brain areas in awake behaving animals while simultaneously recording changes in activity. Multichannel recording systems and powerful computing capabilities will allow study of the spatial properties of oscillations. We can use this technique to specifically disrupt synchrony between regions to assess how activity in different brain areas co-ordinates information.

Conclusions

Considerable evidence now suggests that parvalbumin-containing interneurons may be essential for normal gamma frequency activity, in particular for the long-range synchronisation of network oscillations. Optogenetic techniques seem to provide us with the tools that will enable us to probe the contribution of individual neuronal classes to the generation of gamma, and other, types of network activity implicated in consciousness.

References

- Ashton, H., Reid, K., Marsh, R., Johnson, I., Alter, K. and Griffiths, T. (2007). High frequency localised “hot spots” in temporal lobes of patients with intractable tinnitus: A quantitative electroencephalographic (QEEG) study. *Neuroscience Letters*, 426, 23–28.
- Baldeweg, T., Spence, S., Hirsch, S. R. and Gruzelier, J. (1998). Gamma-band electroencephalographic oscillations in a patient with somatic hallucinations. *Lancet*, 352, 620–621.
- Brady, D. R. and Mufson, E. J. (1997). Parvalbumin-immunoreactive neurons in the hippocampal formation of Alzheimer’s diseased brain. *Neuroscience*, 80, 1113–1125.
- Berger, H. (1929). Über das Elektroenkephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, 87, 527–570.
- Bernstein, H. G., Johnson, M., Perry, R. H., LeBeau, F. E. N., Dobrowolny, H. and Perry, E. K. (2010). Loss of parvalbumin-containing hippocampal interneurons in dementia with Lewy bodies. (in preparation 2010).
- Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G. and Deisseroth, K. (2005). Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neuroscience*, 8, 1263–1268.
- Cardin, J. A., Carlen, M., Meletis, K., Knoblich, U., Zhang, F., Deisseroth, K., Tsai, L. H. and Moore, C. I. (2009). Driving fast-spiking cells induces gamma rhythm and controls sensory responses. *Nature*, 459, 663–667.
- Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J. and Sergent, C. (2006). Conscious, pre-conscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences*, 10, 204–211.
- Engel, A. K., Fries, P., König, P., Brecht, M. and Singer, W. (1999). Temporal binding, binocular rivalry, and consciousness. *Conscious Cognition*, 8, 128–151.
- Fisahn, A., Pike, F. G., Buhl, E. H. and Paulsen, O. (1998). Cholinergic induction of network oscillations at 40 Hz in the hippocampus in vitro. *Nature*, 394, 186–189.
- Freund, T. F. and Buzsáki, G. (1996). Interneurons of the hippocampus. *Hippocampus*, 6, 347–470.
- Fries, P., Reynolds, J. H., Rorie, A. E. and Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*, 291, 1560–1563.
- Fries, P., Roelfsema, P. R., Engel, A. K., König, P. and Singer, W. (1997). Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry. *Proceedings of the National Academy of Sciences of the USA*, 94, 12699–12704.

- Grillner, S., Markram, H., De Schutter, E., Silberberg, G. and LeBeau, F. E. N. (2005). Microcircuits in action—from CPGs to neocortex. *Trends in Neurosciences*, 28, 525–533.
- Han, X. and Boyden, E. S. (2007). Multiple-color optical activation, silencing, and desynchronization of neural activity, with single-spike temporal resolution. *PLoS One*, 2, e299.
- Klausberger, T., Magill, P. J., Marton, L. F., Roberts, J. D., Cobden, P. M., Buzsáki, G. and Somogyi, P. (2003). Brain-state- and cell-type-specific firing of hippocampal interneurons in vivo. *Nature*, 421, 844–848.
- LeBeau, F. E. N., Traub, R. D., Monyer, H., Whittington, M. A. and Buhl, E. H. (2003). The role of electrical signaling via gap junctions in the generation of fast network oscillations. *Brain Research Bulletin*, 62, 3–13.
- Lewis, D. A. and Gonzalez-Burgos, G. (2008). Neuroplasticity of neocortical circuits in schizophrenia. *Neuropsychopharmacology*, 33, 141–165.
- Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W. and Rodriguez, E. (2007). Synchronization of neural activity across cortical areas correlates with conscious perception. *Journal of Neuroscience*, 27, 2858–2865.
- Roopun, A. K., Cunningham, M. O., Racca, C., Alter, K., Traub, R. D. and Whittington, M. A. (2008). Region-specific changes in gamma and beta2 rhythms in NMDA receptor dysfunction models of schizophrenia. *Schizophrenia Bulletin*, 34, 962–973.
- Sederberg, P. B., Kahana, M. J., Howard, M. W., Donner, E. J. and Madsen, J. R. (2003). Theta and gamma oscillations during encoding predict subsequent recall. *Journal of Neuroscience*, 23, 10809–10814.
- Singer, W. and Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Reviews in Neuroscience*, 18, 555–586.
- Sohal, V. S., Zhang, F., Yizhar, O. and Deisseroth, K. (2009). Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. *Nature*, 459, 698–702.
- Somogyi, P. and Klausberger, T. (2005). Defined types of cortical interneurone structure space and spike timing in the hippocampus. *Journal of Physiology*, 562, 9–26.
- Spencer, K. M., Nestor, P. G., Niznikiewicz, M. A., Salisbury, D. F., Shenton, M. E. and McCarley, R. W. (2003). Abnormal neural synchrony in schizophrenia. *Journal of Neuroscience*, 23, 7407–7411.
- Tononi, G. and Koch, C. (2008). The neural correlates of consciousness: An update. *Annals of the New York Academy of Science*, 1124, 239–261.
- Traub, R. D., Bibbig, A., LeBeau, F. E. N., Buhl, E. H. and Whittington, M. A. (2004). Cellular mechanisms of neuronal population oscillations in the hippocampus in vitro. *Annual Reviews in Neuroscience*, 27, 247–278.
- Uhlhaas, P. J. and Singer, W. (2006). Neural synchrony in brain disorders: Relevance for cognitive dysfunctions and pathophysiology. *Neuron*, 52, 155–168.
- Whittington, M. A., Traub, R. D. and Jefferys, J. G. (1995). Synchronized oscillations in interneuron networks driven by metabotropic glutamate receptor activation. *Nature*, 373, 612–615.

Dopamine modulation of decision making processes

Anthony A. Grace

The dopamine system has a number of functions in the brain, from regulation of movement to controlling emotional expression and facilitating cognitive or conscious processes. Dopamine exerts these actions via its regulation of systems interactions. In particular, the influence of this transmitter system on the balance of afferent input to the nucleus accumbens potentially modulates decision-making processes. An increase in dopamine outflow potentiates afferents from the hippocampus subiculum, which are involved in context-dependent processes. In contrast, dopamine will also attenuate the ability of the prefrontal cortex to switch behavioral tasks. In this manner, the dopamine system is positioned to alter the propensity to focus conscious attention on a current task versus flexibly altering behavior to more effectively achieve goals.

The dopamine (DA) system has a number of roles in the regulation of the brain. Among the first of its functions was elucidated with the discovery that loss of DA in the striatum resulted in the movement disorder Parkinson's disease (Hornykiewicz, 1962). However, a more complex role in cognitive function was revealed on the discovery that antipsychotic drugs treat schizophrenia primarily by blocking dopamine receptors in the limbic, or emotional, areas of the brain (Carlsson and Lindqvist, 1963). Thus, in schizophrenia, the dopamine system is dysregulated in such a manner as to cause hyper-release of DA within the limbic/associative striatum (Laruelle and Abi-Dargham, 1999). The role of DA has expanded since this time, particularly given data showing that it is involved in working memory formation in the prefrontal cortex of the brain; the highest center of cognitive integration in humans. DA also plays an important role in emotional regulation, since the DA input to emotional areas such as the amygdala are known to influence emotional learning processes (Rosenkranz and Grace, 2002). Therefore, it is clear that DA plays a central role in a wide variety of brain functions. In this review, I will focus primarily on the impact of DA in regulating motivated behavior, and how this can influence conscious thought processes.

One region of the brain in which the DA system has a powerful influence, and which is central to a number of behavioral functions ranging from reward to motivation, is the ventral striatum. The ventral striatum, including the nucleus accumbens, is a brain region rich in interconnections with areas involved in higher cognitive function, motivation, and reward. Although not an area directly involved in cognitive or conscious thought processes, its connectivity bestows this area with the capacity to regulate a number of aspects of cognitive control and goal-directed behavior. Indeed, individual neurons within the ventral striatum receive inputs from the amygdala, the ventral hippocampus, and the prefrontal cortex (O'Donnell and Grace, 1995); all regions involved in emotional and cognitive regulation. Although each of these afferents uses glutamate as a transmitter, the functional impact of these systems on the ventral striatal system is markedly different. Moreover, the DA system provides a potent influence over functional interactions of these afferents within the ventral striatum, enabling a shift between dominance depending on the state of the organism. The ventral striatum, in turn, provides potent outputs that affect the DA system, as well as the thalamic nucleus that interconnects with the prefrontal cortex. Therefore, these functional loops are positioned to alter the effective weight of each of these systems in regulating behavior. In particular, this system has potent regulatory influences on reward, attention, and the control of goal-directed behavior, and its dysfunction leads to disorders ranging from drug abuse to schizophrenia and depression. How the ventral striatum and its afferents are involved in such complex tasks can be better understood by examining the function of each of these systems, and how they compete for influence within the ventral striatum.

Ventral subiculum of the hippocampus

The hippocampus is organized into a number of subcomponents. The entorhinal cortex is the region that receives inputs from several regions, and the ventral subiculum (functionally equivalent to the anterior hippocampus in primates) is generally considered to be the output of this region (O'Mara, 2005). However, each of the subcomponents in turn has specialized functions. For example, the dorsal (posterior) sections of the hippocampus are known to provide information regarding location in space, and the neurons receive afferent input that enables them to show properties of "place cells" – i.e., neurons that fire when an organism is located in a particular location in space (O'Keefe, 1979). As one moves ventrally through the rat hippocampus, these place columns receive inputs from various aspects of the limbic system, modifying the "place" information to contain data related to emotion. At its ventralmost output region, the ventral subiculum, the

information content has been expanded to include emotion, which I have proposed to be the emotional relevance of a location in space (Grace, 2010). The behavioral and emotional significance of a location could be interpreted as context. Indeed, it has been proposed that the ventral subiculum is the part of the brain in which context is represented (Sharp, 1999), and as such plays a role in context-specific interpretations of stimuli.

The ventral subiculum has a number of outputs of relevance for cognitive control and motivation (O'Mara, 2005). It has a prominent output to the prefrontal cortex, which mediates aspects of memory formation. Memory-guided behavior is generally believed to occur along several dimensions; among these are retrospective and prospective information-driven behaviors. Episodic memory is an example in which information is gleaned from past events to form memories. This episodic memory can then be used in a prospective manner to plan for future strategies. Thus, if we are walking along our normal route and a bridge is out, we can use this retrospective memory to recall alternate routes. However, the following day, we can recall our past experience with the bridge in order to form a prospective plan to avoid that route. Our studies have shown that retrospective memory is dependent on the hippocampal formation (Goto and Grace, 2008). However, for this episodic memory to be utilized in a prospective, predictive manner, we found that information must be transferred to the prefrontal cortex in a manner that depends on both D1 and D2 dopamine receptors in the frontal cortical region. In this way, memory of events can be used to guide future behavior.

In addition to providing this influence over memory events, the ventral subiculum also plays a prominent role in regulating information flow in a context-dependent manner subcortically. A prominent output of the subiculum is to the ventral striatum/nucleus accumbens region. Here, the ventral subiculum innervates spines of medium spiny GABAergic neurons. These medium spiny neurons are the crossroads of a number of afferent inputs. Thus, as cited above, it has been shown that single accumbens medium spiny neurons receive excitatory glutamatergic input from at least three sources: the prefrontal cortex, the nucleus accumbens, and the amygdala (O'Donnell and Grace, 1995). However, despite the fact that all of these afferents contain the same transmitter, the impact of the ventral subicular input on the accumbens is quite different. Thus, the subiculum is found to produce long-duration (hundreds of millisecond) depolarized plateau "up" states in accumbens neurons (O'Donnell and Grace, 1995). These long-duration up states are due to properties of the afferent input and the membrane properties of the medium spiny neurons. In the absence of this afferent input, the membrane potential of the accumbens neuron is very hyperpolarized, and is not likely to generate spike activity (and hence information flow) in response to even strong

inputs from other regions. However, when the ventral subiculum drives the accumbens into an up state, the membrane potential is sufficiently depolarized to allow the neuron to generate spikes to comparatively small excitatory afferent inputs. Thus, the ventral subiculum “gates” information flow through the nucleus accumbens (O’Donnell and Grace, 1995). This could be of particular relevance for inputs from areas such as the prefrontal cortex. Thus, the prefrontal cortex would depend on an input from the hippocampus in order to generate spike activity. This accumbens activity would then be transmitted via the ventral pallidum to the mediodorsal nucleus of the thalamus, which itself innervates the prefrontal cortex (Grace, 2000). Thus, information flow along the prefrontal-accumbens-pallidal-thalamic-prefrontal loop would be gated by ventral subicular inputs to the accumbens (Figure 1). Given that the subiculum activity is context-dependent, then the subiculum would be capable of filtering prefrontal input based on its relevance to the current context. If the input from the prefrontal cortex is relevant to the context, this information is passed along the loop and reinforced via a feedback to the prefrontal cortex. If however the information is out of context, information flow would be blocked. Thus, the hippocampus will keep the organism focused on a task, while blocking extraneous influences that attempt to divert the individual from performing the task at hand. The input coming from the hippocampus subiculum is further modulated by the dopamine system. Thus, activation of dopamine D1 receptors will potentiate subicular inputs to the accumbens (Goto and Grace, 2005b).

Prefrontal cortex and behavioral flexibility

The prefrontal cortex is the highest centre of integration in the mammalian brain. As such, it has influences over a number of brain activities. In fact, if one were to catalogue the influences of the prefrontal cortex, it would seem that there was precious little else for the rest of the brain to do, as this would involve: executive function, working memory, prospective memory, emotional regulation, alternation of behavior, delay processes, sensory integration, motor control, etc. How could one region perform all of these functions? In my opinion, this would not be feasible, and indeed would cause the organism to be rather vulnerable to damage to this area. Indeed, animals in which the prefrontal cortex has not as yet developed can still perform many of these functions. Furthermore, it is clear that lesion of this area does not totally incapacitate all animal or human behaviors, as this was once a common treatment for severe mental disorders. One resolution to this dilemma is if the prefrontal cortex is involved in the *regulation* of each of these functions. Thus, the prefrontal cortex, via its afferents to the amygdala, can

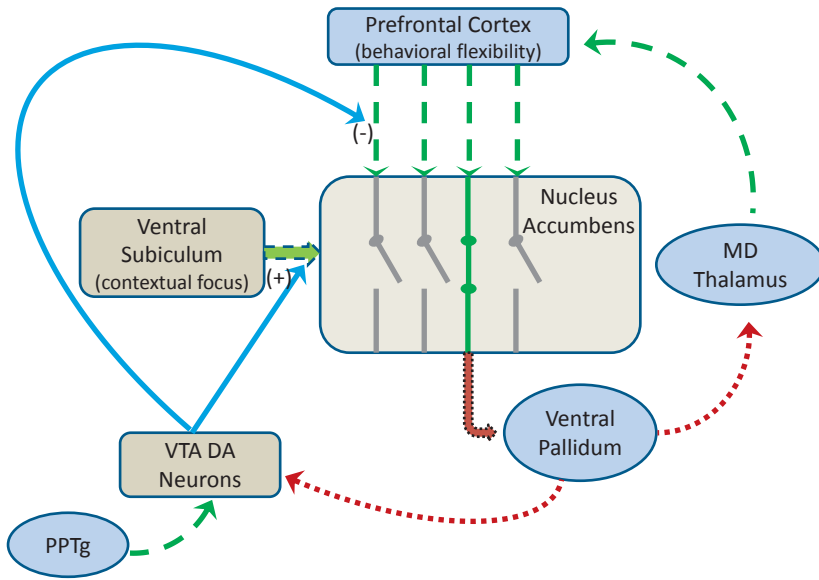


Figure 1. The ventral subiculum of the hippocampus is a structure whose activity is determined by the environmental context of the organism, and is presumed to gate information flow according to context. Thus, the subiculum drives a context-specific subset of neurons in the nucleus accumbens into up states; only those neurons in the up state will fire spikes when receiving input. In this example, the prefrontal cortex only triggers action potential firing in the neuron in the up state (green lines). This neuron then passes information to the ventral pallidum and mediodorsal (MD) thalamic nucleus and back up to the prefrontal cortex, reinforcing this loop while blocking activity in pathways related to other contexts (gray lines). This circuit also influences DA neuron firing; when the subiculum is activated, the activation of the accumbens removes ventral pallidal inhibition of ventral tegmental area (VTA) DA neurons, causing more of them to fire. The firing neurons can then be driven to burst fire by the pedunculopontine tegmentum (PPTg). The increase in DA transmission in the accumbens further potentiates ventral subicular input while attenuating prefrontal cortical input, shifting the balance of information flow in this region. Short dashed arrows = inhibitory; Long dashed arrows = excitatory; Solid grey arrows = modulatory.

effectively suppress amygdala responses to aversive stimuli as well as emotional learning (Rosenkranz and Grace, 2002). Via its afferents to the accumbens, the prefrontal cortex can regulate drug responses and supply an afferent bias to the hippocampal input, as well as compete with this input (see below). Therefore, the prefrontal cortex is positioned not to mediate all of these behaviors, but instead to provide a regulatory “top-down” modulation of these behaviors.

One of the behaviors that the prefrontal cortex is found to potently influence is behavioral flexibility (Clarke et al., 2004). Thus, when an animal is performing a task that leads to a reward, the prefrontal cortex does not need to be involved. However, if the task is found to be ineffective in producing a reward, the prefrontal cortex can activate to enable the organism to select a different strategy. This is made clear in tasks such as the Wisconsin card-sort test. In this task, a person is asked to sort cards according to one dimension of information – e.g., the shape of the objects on the card. Such a task does not activate the prefrontal cortex. However, if the criterion for a correct response now changes such that the old strategy is now ineffective (e.g., now the number of objects on the card is the rewarded criterion), the prefrontal cortex activates, enabling the person to ignore the previously reinforced behavior and to select a novel strategy (Berman et al., 1995). If on the other hand, the prefrontal cortex is damaged, the organism will show perseveration – i.e., it will continue to perform the task according to the previous criteria, despite the lack of reinforcement (Ragozzino et al., 2003).

As reviewed above, the prefrontal cortex also sends glutamatergic inputs to the nucleus accumbens. Moreover, the prefrontal cortical inputs to the accumbens, like those of the subiculum, are also modulated by DA. However, this modulation occurs via a distinct DA receptor subtype; i.e., the D2 receptor family. Moreover, DA is found to attenuate input coming from the prefrontal cortex (Goto and Grace, 2005b; O'Donnell and Grace, 1994). Thus, the DA system affects the subiculum and prefrontal cortical inputs to the accumbens in opposite manners: it potentiates the hippocampal input via an action on D1 receptors while attenuating the prefrontal cortical input via an action on D2 receptors.

Dopamine system regulation in the control of accumbens afferents

The DA system has its origin within the midbrain, in a cluster of neurons within the ventral tegmental area and substantia nigra. Although the number of DA neurons in this region is comparatively limited compared to many other brain nuclei, their widespread connections and extensive collateralization cause this neuron class to have a substantial impact on its postsynaptic targets. Indeed, it is estimated in the rat that each DA neuron gives rise to over 1,000,000 terminals (Anden et al., 1966). Given that these projections are furthermore topographically organized, this would enable small clusters of DA neurons to exert a substantial impact on their target neurons.

The DA neurons themselves are rather unusual in their activity states. Thus, DA neurons typically exist across three activity states. First, a large proportion of the DA neurons in the baseline, unstimulated state are quiescent and non-firing;

this is estimated to be approximately 50% of the population of neurons (Grace and Bunney, 1984b). DA neurons in this state may be considered “in reserve,” and can be activated when the system is placed under demand. This state is believed to be maintained by a potent inhibitory input arising from the ventral pallidum (Floresco et al., 2003). When this inhibition is released, DA neuron activity is driven via an endogenous pacemaker conductance. This pacemaker conductance drives DA neurons in a pacemaker-like pattern when there is no afferent drive, as is observed *in vitro* in brain slices (Grace and Onn, 1989); *in vivo*, the pattern is one of a slow, irregular discharge activity. This would be the “idling” state of active DA neurons. The proportion of DA neurons that are firing spontaneously is termed the population activity. The third state occurs in response to a behaviorally relevant stimulus; this causes the DA neurons to emit a burst of action potentials (Grace and Bunney, 1984a). This burst is driven by glutamate acting on NMDA receptors (Chergui et al., 1993), and is proposed to be the behaviorally salient output of the DA system. Thus, when an organism encounters a behaviorally salient stimulus, such as one predicting reward, the DA system is caused to burst fire (Schultz, 1998). The glutamatergic input that drives burst firing in the limbic DA neurons arises primarily from a brainstem region known as the pedunculopontine tegmentum (PPTg; Floresco et al., 2003; Lodge and Grace, 2006). DA neurons will burst fire only upon NMDA receptor drive. However, in order to activate NMDA receptors, the neuron must be in a depolarized state; otherwise, there is a magnesium block of the NMDA channel (Mayer et al., 1984). As a consequence, only DA neurons that are spontaneously active will respond to PPTg input with a burst of spikes.

As mentioned, the population activity of the DA neurons (i.e., the proportion of DA neurons firing spontaneously) is regulated by a potent GABAergic input arising from the ventral pallidum. The ventral pallidum, in turn, is inhibited by the ventral striatum/nucleus accumbens. Furthermore, the ventral striatum can be potently excited by the ventral subiculum of the hippocampus. Thus, the subiculum is positioned to control the population activity of the mesolimbic DA system. Indeed, our studies show that activation of the ventral subiculum will indeed increase the DA neuron population activity by more than 2 fold via the subiculum-accumbens-ventral pallidal pathway (Floresco et al., 2003). This will have an important consequence for the state of responsiveness of the DA system. Thus, a behaviorally salient stimulus will activate burst firing of DA neurons via an activation of the PPTg-glutamate input. However, it will only drive burst firing in the population of DA neurons that is already firing spontaneously; a factor that is regulated by the subiculum. Thus, the PPTg will supply the behaviorally salient activation, but the amplitude of this activation will depend on the number of DA neurons firing, which is driven by the ventral subiculum. The ventral

subiculum, therefore, provides the level of amplification of the signal (Lodge and Grace, 2006) (Figure 2). This is consistent with the role of the subiculum in mediating context-dependent processes. Thus, if one is in a benign environment where there is little expectation of a serious threat or probability of action leading to a large reward, the subiculum would keep the DA system in a baseline state. If a signal occurs, such as a noise in the background, the activation of the DA system would be minimal, and the urgency with which a response is generated is low. If, however, one is in an environment in which there is a high expectation of reward (e.g., hunting) or danger (e.g., a war zone), the subiculum would drive the DA system at a high level, causing the stimulus to trigger burst firing in a large number of DA neurons and attention to be strongly drawn toward the object in deference to other ongoing activity. Therefore, the contingencies of the environmental context would, via the ventral subiculum, set the response level of the DA system appropriately.

When there is a system such as the one described, there is the potential that disruption of the system could lead to altered behavior. Indeed, we have found that dysfunction of DA neuron population activity can play a substantial role in disease states. Thus, in schizophrenia patients, imaging studies have shown the presence of hyperactivity within the anterior hippocampus (Malaspina et al., 1999). Moreover, it is known that in schizophrenia, the DA system is hyper-responsive, in that a stimulus such as amphetamine will cause a greater DA release (in terms of raclopride displacement) than in a normal subject (Laruelle and Abi-Dargham, 1999). In a developmental animal model of schizophrenia (Moore et al., 2001), we have found that the ventral subiculum is indeed hyperactive (Lodge and Grace, 2007), possibly owing to a decreased interneuron function (Lodge et al., 2009). Moreover, the DA neuron population activity is maximally increased at baseline under this condition (Lodge and Grace, 2007). Thus, the “gain” is set to maximum, under all conditions. As a consequence, all stimuli that occur, whether they are behaviorally salient or not, will cause a maximal DA release and draw attention to what has been interpreted as a highly threatening stimulus. Thus, the normally benign stimulus would be given an abnormally high level of salience. Such “aberrant salience” has been proposed to lead to delusional states (Kapur, 2003), possibly due to this abnormally amplified DA signal. In other words, even if the object itself is benign, the attribution of a high level of motivational salience to the object would lead to construction of a condition in which the object does have the level of salience assigned to it, even if this construct is not based on fact. Indeed, inactivation of the ventral subiculum in this rat model was found to reverse both the increase in DA neuron population activity as well as the hyper-responsivity to amphetamine (Lodge and Grace, 2008).

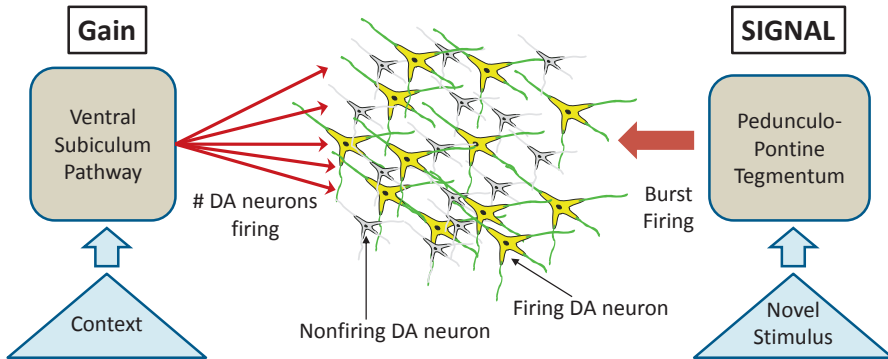


Figure 2. Activity of the DA neurons is determined by two factors: population activity (i.e., the number of neurons firing) and burst firing. The pathway originating in the ventral subiculum determines the number of DA neurons firing; the higher the subicular activity, the more neurons that are active. In contrast, the pedunclopontine tegmentum causes spontaneously firing DA neurons to emit bursts of spikes. When a novel stimulus is detected, the pedunclopontine tegmentum activates, causing spontaneously firing DA neurons to emit a burst of spikes. The number of DA neurons that are driven to burst fire is determined by the ventral subiculum as a function of the environmental context. In more activating contexts, the subiculum causes a greater number of DA neurons to fire spontaneously, and hence a larger proportion are driven to burst fire by the pedunclopontine tegmentum. Thus, the amplitude of the response to the signal is determined by the gain provided by the ventral subiculum.

Prefrontal cortical-hippocampal interactions and goal-directed behavior

The interaction of the prefrontal cortex, the subiculum and the DA system is also proposed to play a substantial role in guiding behavioral responses and decision-making. Thus, as reviewed above, afferents from the subiculum and prefrontal cortex will drive accumbens neurons via a glutamatergic process. Moreover, the DA system regulates this drive in different manners, with DA D1 stimulation potentiating the ventral subicular input, while DA D2 receptor stimulation will attenuate prefrontal cortical drive. However, the impact of these afferents was also found to be potently influenced by synaptic plasticity. Thus, high-frequency theta burst stimulation of the hippocampal input to the accumbens is known to produce a long-term potentiation (LTP) of inputs from this region. In other words, a strong drive of the hippocampal afferent will potentiate further input coming from this region. The same condition will also occur if the prefrontal cortex is

stimulated; thus, high-frequency stimulation of the prefrontal cortex will induce LTP on this afferent system. However, there is also another face to this response: high-frequency stimulation of the subiculum will induce LTP in this afferent, but will cause long-term depression (LTD) of the prefrontal input. Conversely, high-frequency stimulation of the prefrontal input will cause LTP in the stimulated input and LTD in the subicular input (Goto and Grace, 2005a). Thus, the prefrontal cortex and the hippocampus will compete with each other for the control over the accumbens, with the system that is driven the strongest winning. The DA system will provide a further bias to this system as well. Thus, increases in DA transmission will potentiate the hippocampal subicular input while attenuating the prefrontal cortical input, thereby biasing subicular drive over prefrontal drive. Given that activation of the hippocampus will also drive DA neuron activity, the subicular-DA-subicular potentiation loop would be self-perpetuating.

This condition of subiculum-prefrontal competition for control of the accumbens may be interpreted on a functional basis. Although speculative, such a condition could draw a link between DA regulation of synaptic plasticity and goal-directed behavior. I propose that this system is organized in such a way as to control most of ongoing behavior and direction of attention. Suppose that you are engaged in a task, such as having an engaging conversation with a colleague. This is the ongoing activity, defined in the context of interpersonal interactions. As a result, the ventral subiculum would be activated for that context, leading to the induction of LTP in the subiculum-accumbens pathway. If the interaction is engaging, the rewarding nature of the interaction would activate the DA system, further potentiating this engagement while screening out other influences. Your full attention is focused on the person and the conversation, and context-dependent LTP in the subiculum-accumbens pathway is reinforcing this highly focused attention while preventing interference from distractors. But let's say that during this conversation, someone with whom you would like to talk enters the room and calls out your name. Since this is an interaction with which you are highly motivated to have, this would decrease the relative reinforcing value of your current conversation. Thus, just like in the event where lack of an expected reward will inhibit DA neuron firing (Schultz, 1998), this decrease in reinforcing value of the current conversation would decrease the DA input to the accumbens, diminishing the subicular focus on the current task, and through disinhibition allowing activation of the prefrontal cortex. The prefrontal cortex would then initiate "behavioral flexibility," causing you to disengage from your current conversation (by inducing LTD in the subiculum-accumbens pathway and LTP in the prefrontal-accumbens pathway). The prefrontal cortical control would then drive you to an alternate context (i.e., the person that entered the room). Once the new conversation is engaged, the resulting reward would activate the DA system, disengage

the prefrontal drive (via D2 inhibition of prefrontal afferents and LTD in the prefrontal-accumbens pathway) and engage the subiculum (via D1 potentiation of hippocampal afferents and LTP in the hippocampal-accumbens pathway), thus locking attention onto the new conversation. In this way, attention can be focused in a flexible manner on an ongoing activity provided it has some reward value, but also to disengage when either the current activity fails to produce a reward or a competing, more rewarding task presents itself.

This would be the normal manner in which the system would interact – LTP and LTD existing in a flexible, reversible manner to direct and hold attentional states and to switch engagements as the contingencies demand. However, it would also be possible to disrupt such a delicate, effective goal-directed balance. One way this could happen is with drugs of abuse. Drugs have been described as “hijacking” the brain (Bechara, 2005). Our studies revealed an interesting phenomenon that is related to this interaction. Following repeated cocaine treatment, we found an induction of a DA-dependent LTP in the subiculum-accumbens pathway, that leads to perseverative behavior (Goto and Grace, 2005a). Thus, this flexible system that will engage and disengage depending on contingencies is now thwarted, with the cocaine causing the system to focus engagement on the current context – i.e., taking the drug. Under this state, the act of taking the drug becomes the reinforced context via induction of LTP in the subiculum-accumbens pathway. Such a condition would not depend on the drug being reinforcing, since it is now the act of taking the drug that drives the focus of the system. This could be representative of the addicted brain, in which the “liking” of the drug is replaced by “wanting” the drug, independent of the reinforcing properties (Robinson and Berridge, 1993).

Dopamine, cognitive control and consciousness

The contribution of the limbic system to the control of directed attention and motivation is outlined, and a speculative synthesis of how these systems could interact to influence decision-making is presented. The DA system provides an important modulatory influence over these activities. DA is in itself not a reward, but instead is a system that controls balance and level of activation – it is the lubricant that facilitates focus and enables drives to change. The decision-making process in this model is driven in a reward-based manner. But it may be that this reward is what engages us to perform activities. Just as DA is required to allow movement, it is also required to provide motivation and drive. Perhaps it is the system that enables us to “feel” the consequences of our actions as correct (i.e., rewarding) or aberrant, and to direct the systems to engage to either maintain performance of the rewarded task or to change behaviors to test alternate possibilities. Such

systems are likely present at many levels of phylogeny beyond the human. What makes us human may be the ability to become consciously aware of the decision process, to the benefit or the detriment of the actor.

How does this relate to consciousness? Consciousness, as stated many times, is the “hard” problem. But we also know that conscious thought is not a static state, but a dynamic state that is modulated by environment and behavioral necessities. What I have described here is how consciousness can be directed; that the focusing of conscious attention is potently modulated by circuits that evaluate the impact of the directed attention, whether it is appropriate, and provide a circuit-based mechanism by which this conscious state can be biased. The ability to use prospective information to inform a conscious decision, or to shift conscious attention to attain the greatest advantage, are likely systems that have been built evolutionarily from comparatively simple constructs used to react automatically to the environment, to the state in the human in which the flow of conscious thought is organized and directed. When this flow of thought is disrupted, as occurs in schizophrenia, or is misappropriated, as in drug abuse, the fine control of conscious thought can no longer be effectively managed to deal with the processing of complex information in a way that would direct optimal behavior.

References

- Anden, N. E., Fuxe, K., Hamberger, B. and Hokfelt, T. (1966). A quantitative study on the nigro-neostriatal dopamine neuron system in the rat. *Acta Physiologica Scandinavica*, 67, 306–312.
- Bechara, A. (2005). Decision making, impulse control and loss of willpower to resist drugs: A neurocognitive perspective. *Nature Neuroscience*, 8, 1458–1463.
- Berman, K. F., Ostrem, J. L., Randolph, C., Gold, J., Goldberg, T. E., Coppola, R., Carson, R. E., Herscovitch, P. and Weinberger, D. R. (1995). Physiological activation of a cortical network during performance of the Wisconsin Card Sorting Test: A positron emission tomography study. *Neuropsychologia*, 33, 1027–1046.
- Carlsson, A. and Lindqvist, M. (1963). Effect of chlorpromazine or haloperidol on formation of 3-methoxytyramine and normetanephrine in mouse brain. *Acta Pharmacologica et Toxicologica*, 20, 140–144.
- Chergui, K., Charlety, P. J., Akaoka, H., Saunier, C. F., Brunet, J. L., Buda, M., Svensson, T. H. and Chouvet, G. (1993). Tonic activation of NMDA receptors causes spontaneous burst discharge of rat midbrain dopamine neurons in vivo. *European Journal of Neuroscience*, 5, 137–144.
- Clarke, H. F., Dalley, J. W., Crofts, H. S., Robbins, T. W. and Roberts, A. C. (2004). Cognitive inflexibility after prefrontal serotonin depletion. *Science*, 304, 878–880.
- Floresco, S. B., West, A. R., Ash, B., Moore, H. and Grace, A. A. (2003). Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nature Neuroscience*, 6, 968–973.

- Goto, Y. and Grace, A. A. (2005a). Dopamine-dependent interactions between limbic and prefrontal cortical plasticity in the nucleus accumbens: Disruption by cocaine sensitization. *Neuron*, 47, 255–266.
- Goto, Y. and Grace, A. A. (2005b). Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nature Neuroscience*, 8, 805–812.
- Goto, Y. and Grace, A. A. (2008). Dopamine modulation of hippocampal-prefrontal cortical interaction drives memory-guided behavior. *Cerebral Cortex*, 18, 1407–1414.
- Grace, A. A. (2000). Gating of information flow within the limbic system and the pathophysiology of schizophrenia. *Brain Research – Brain Research Reviews*, 31, 330–341.
- Grace, A. A. (2010). Dopamine system dysregulation as the common pathophysiological basis for schizophrenia psychosis, psychostimulant abuse, and stress. *Neurotoxicology Research*, (in press).
- Grace, A. A. and Bunney, B. S. (1984a). The control of firing pattern in nigral dopamine neurons: Burst firing. *Journal of Neuroscience*, 4, 2877–2890.
- Grace, A. A. and Bunney, B. S. (1984b). The control of firing pattern in nigral dopamine neurons: Single spike firing. *Journal of Neuroscience*, 4, 2866–2876.
- Grace, A. A. and Onn, S. P. (1989). Morphology and electrophysiological properties of immunocytochemically identified rat dopamine neurons recorded in vitro. *Journal of Neuroscience*, 9, 3463–3481.
- Hornykiewicz, O. (1962). [Dopamine (3-hydroxytyramine) in the central nervous system and its relation to the Parkinson syndrome in man.]. *Dtsch Med Wochenschr*, 87, 1807–1810.
- Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry*, 160, 13–23.
- Laruelle, M. and Abi-Dargham, A. (1999). Dopamine as the wind of psychotic fire: New evidence from brain imaging studies. *Journal of Psychopharmacology*, 13, 358–371.
- Lodge, D. J., Behrens, M. M. and Grace, A. A. (2009). A loss of parvalbumin-containing interneurons is associated with diminished oscillatory activity in an animal model of schizophrenia. *Journal of Neuroscience*, 29, 2344–2354.
- Lodge, D. J. and Grace, A. A. (2006). The hippocampus modulates dopamine neuron responsiveness by regulating the intensity of phasic neuron activation. *Neuropsychopharmacology*, 31, 1356–1361.
- Lodge, D. J. and Grace, A. A. (2007). Aberrant hippocampal activity underlies the dopamine dysregulation in an animal model of schizophrenia. *Journal of Neuroscience*, 27, 11424–11430.
- Lodge, D. J. and Grace, A. A. (2008). Amphetamine activation of hippocampal drive of mesolimbic dopamine neurons: a mechanism of behavioral sensitization. *Journal of Neuroscience*, 28, 7876–7882.
- Malaspina, D., Storer, S., Furman, V., Esser, P., Printz, D., Berman, A., Lignelli, A., Gorman, J. and Van Heertum, R. (1999). SPECT study of visual fixation in schizophrenia and comparison subjects. *Biological Psychiatry*, 46, 89–93.
- Mayer, M. L., Westbrook, G. L. and Guthrie, P. B. (1984). Voltage-dependent block by Mg^{2+} of NMDA responses in spinal cord neurones. *Nature*, 309, 261–263.
- Moore, H., Ghajarnia, M., Geyer, M., Jentsch, J. D. and Grace, A. A. (2001). Selective disruption of prefrontal and limbic corticostriatal circuits by prenatal exposure to the DNA methylation agent methylazoxymethanol acetate (MAM): Anatomical, neurophysiological and behavioral studies. *Schizophrenia Research*, 49(S1–2), 48.

- O'Donnell, P. and Grace, A. A. (1994). Tonic D2-mediated attenuation of cortical excitation in nucleus accumbens neurons recorded in vitro. *Brain Research*, 634, 105–112.
- O'Donnell, P. and Grace, A. A. (1995). Synaptic interactions among excitatory afferents to nucleus accumbens neurons: hippocampal gating of prefrontal cortical input. *Journal of Neuroscience*, 15, 3622–3639.
- O'Keefe, J. (1979). A review of the hippocampal place cells. *Progress in Neurobiology*, 13, 419–439.
- O'Mara, S. (2005). The subiculum: What it does, what it might do, and what neuroanatomy has yet to tell us. *Journal of Anatomy*, 207, 271–282.
- Ragozzino, M. E., Kim, J., Hassert, D., Minniti, N. and Kiang, C. (2003). The contribution of the rat prelimbic-infralimbic areas to different forms of task switching. *Behavioural Neuroscience*, 117, 1054–1065.
- Robinson, T. E. and Berridge, K. C. (1993). The neural basis of drug craving: An incentive-sensitization theory of addiction. *Brain Research – Brain Research Reviews*, 18, 247–291.
- Rosenkranz, J. A. and Grace, A. A. (2002). Dopamine-mediated modulation of odour-evoked amygdala potentials during pavlovian conditioning. *Nature*, 417, 282–287.
- Schultz, W. (1998). The phasic reward signal of primate dopamine neurons. *Advances in Pharmacology*, 42, 686–690.
- Sharp, P. E. (1999). Complimentary roles for hippocampal versus subicular/entorhinal place cells in coding place, context, and events. *Hippocampus*, 9, 432–443.

Undercurrents of consciousness

The endocannabinoid system

Heather Ashton

The endocannabinoid system acts as a neuromodulator, interacting with other neuromodulators such as the opioids. It influences all vital body functions and sets the tone of every conscious thought and feeling. It operates largely below conscious levels but occasionally floats into consciousness where it is perceived as perfect bliss and pain-free calm.

Worlds on worlds are rolling ever
From creation to decay
Like the bubbles on a river
Sparkling, bursting, borne away
P. B. Shelley (1792–1822)

With its 100 billion neurones, each connecting with over 10,000 others and changing activity as often as 200 times a second, the brain is a veritable factory of information-processing which mostly occurs at an unconscious, subconscious or preconscious level. Above the ever-active undercurrents, conscious awareness erupts from time to time like an evanescent bubble representing only a minute proportion of the total information processing. At any one moment we are generally aware of a single train of thought, sensation or emotion which constantly changes in intensity or direction or is replaced by another.

At other times, the bright bubble fades as we become less vividly aware of our thoughts and surroundings, which descend into vagueness and finally disappear as we sink into dreamless sleep – often to rise briefly in modified form during dreaming. Thus the level of consciousness is on a constantly shifting spectrum with attentive awareness at one extreme and dreamless sleep at the other. But the whole brain never sleeps and the undercurrents flow on, unnoticed but powerful.

These levels of consciousness have physiological underpinnings. Neurones communicate via neurotransmitters and neuromodulators. As described by Perry and Young (2002), the chief “executive” neurotransmitters determining the general level of cortical excitability are GABA (gamma aminobutyric acid) and glutamate. Interacting with these primary control mechanisms is a complex series of modulatory transmitter systems, including acetylcholine and the monoamines dopamine, noradrenaline, serotonin, histamine and others. Most of these originate in

discrete subcortical nuclei but have widespread cortical projections which target receptors on many other neurones. A further series of diffusely distributed linked modulatory systems include neuropeptides such as opioids, adenosine, the gas nitric oxide and the endocannabinoid system – the subject of this chapter.

The endocannabinoid system

“I think big decisions are often made – or drafted – over long periods in the unconscious mind where evidence accumulates and arguments are weighed without the conscious man knowing what is going on. In time, everything is ready for the flick of the switch, and it is only of that final flick that we are conscious.”

Matthew Parris, *Times2*, Sept. 27, 2002, p. 5

Possibly the most powerful and ancient of the modulatory systems that influence every conscious thought and feeling is the endocannabinoid system which has been preserved through eons of evolution. Cannabinoid receptors are found in invertebrates such as leeches, molluscs and insects that diverged from vertebrates over 300 million years ago, and are present in all vertebrates. This evolutionary conservation suggests that they serve important physiological functions.

The endocannabinoid system consists of specific cannabinoid receptors (CB₁, CB₂ and probably others), their natural ligands (anandamides and others) and specific enzymes for their biosynthesis and inactivation. In man, the system modulates almost all vital functions, including those clearly involved in consciousness (cognition, learning, memory, perception, mood, sleep, pain, appetite, reward, motivation) and many that do not normally reach consciousness (motor control, cardiovascular regulation, endocrine activity, metabolism, immune reactions). Furthermore, the system interacts closely with other modulator systems and influences the activity of all the neurotransmitters including GABA and glutamate.

Receptors. CB₁ receptors are widely distributed in the brain including the cerebral cortex, especially frontal cortex, limbic system including hippocampus and amygdala, all sensory areas, motor areas including thalamus and central pain pathways, basal ganglia and cerebellum, hypothalamus, pons and medulla. There are more CB₁ receptors in the brain than all the dopamine, noradrenaline and serotonin receptors combined and ten times more than opioid receptors. CB₁ receptors are also present in the spinal cord and peripheral nerve terminals in the gut and fat cells (adipocytes).

CB₂ receptors are distributed mainly in immune tissues and inflammatory cells, including spleen, tonsils, thymus, lymphocytes, mononuclear cells and macrophages, and there are also some in the brain.

Both CB₁ and CB₂ receptors are metabotropic receptors coupled to a second messenger G_{i/o} protein. Their mechanisms of action are shown in Figure 1. The final result of their activation is inhibition of neuronal depolarisation, decreased action potential generation, decreased release of neurotransmitters, either excitatory or inhibitory, and hence reduced impulse propagation. Other receptors, such as for monoamines or opioids, on the same post-synaptic neurone as CB₁ receptors may share common mechanisms, thus setting the scene for interactions between the different modulator systems (Figure 1).

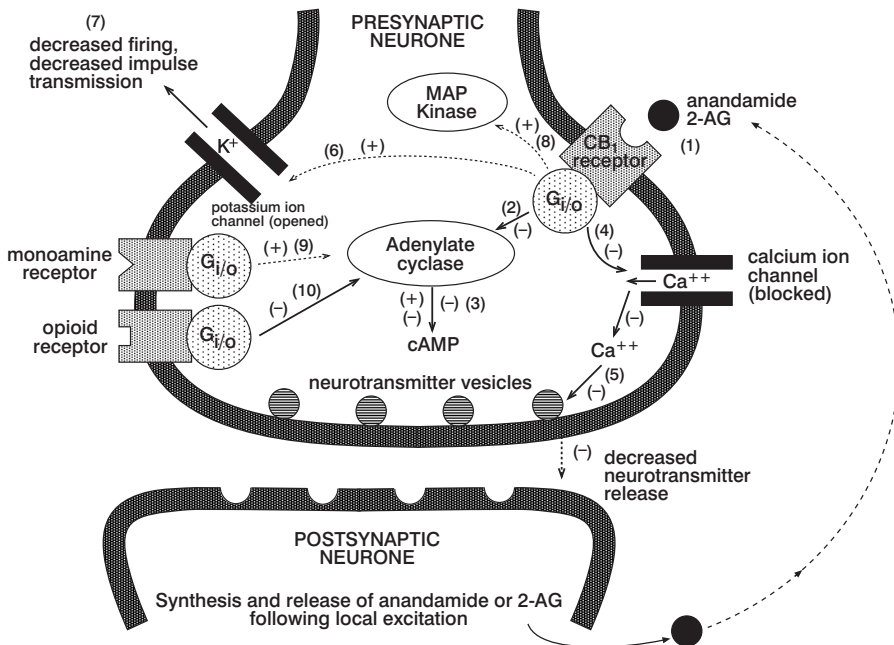


Figure 1. Signal transduction mechanisms mediated by CB₁ receptors.

Local excitation of a post-synaptic neurone triggers the synthesis of anandamide or 2-AG which are released into the extracellular space and act retrogradely as agonists on presynaptic CB₁ receptors. The CB₁ receptor (1) is coupled to a second messenger G_{i/o} protein. Via this protein, activation of the receptor inhibits the enzyme adenylate cyclase (2) and decreases the production of cAMP (3). Via the G-protein, the inward flow of calcium ions is blocked (4), decreasing the release of neurotransmitters (5). Also via the G-protein, the outward flow of potassium ions is enhanced (6), resulting in decreased neuronal firing and decreased impulse transmission (7). Stimulation of the G-protein also activates MAP kinase (8), affecting intracellular gene expression. Other receptors on the same neurone (for monoamines and/or opioids) may activate their own G-proteins but share a common adenylate cyclase, which they may stimulate (9) or inhibit (10). (Christie and Vaughan, 2001; Alger, 2004).

Endocannabinoids. The endogenous ligands for both CB₁ and CB₂ receptors are a family of arachidonic acid (fatty acid) derivatives. The first to be discovered were anandamide (N-arachidonoyl-ethanolamine, AEA) and 2-arachidonoyl-glycerol (2-AG), but three more cannabinoid receptor agonists have recently been proposed (De Petrocellis et al., 2009). Endocannabinoids are present in the brain in the same areas as CB₁ receptors. Unlike monoamine neurotransmitters, anandamide and 2-AG are not stored in vesicles but are synthesised in post-synaptic neuronal membranes and released at discrete loci ‘on demand’ following physiological stimuli such as neuronal depolarisation. After release from the post-synaptic membrane, they act retrogradely as agonists on presynaptic CB₁ receptors (See Figure 1). However, they are rapidly inactivated by enzymatic hydrolysis and removed from the site of action by neuronal reuptake mechanisms. Details of the biosynthesis, release and inactivation of anandamide and other endocannabinoids are given by De Petrocellis et al. (2009). The localised action and rapid inactivation of anandamide and 2-AG is in sharp contrast to the actions of the cannabis plant-derived cannabinoid THC (Δ^9 -tetrahydrocannabinol) which is widely distributed, stimulating simultaneously CB₁ and CB₂ receptors in all areas, and is very slowly eliminated, producing long-lasting and abnormal physiological effects. The effects of THC, especially in high doses, override the delicate balance within the endocannabinoid system and cause severe cognitive impairment, sensory distortion, and perturbations of consciousness.

Endocannabinoid tone

“I have the growing belief that the endocannabinoid system ...
constitutes a major modulator/messenger supersystem,
which is at least as important as the monoaminergic, purinergic
and cholinergic systems. Furthermore, these modulator
systems work hand in hand”

Attila Köfalvi in *Cannabinoids and the Brain* (2008)

How are you feeling today? Happy/depressed? Relaxed/anxious? Pain-free/aching? A great many conscious feelings are determined by the underlying, and unconscious, “tone” of the endocannabinoid system – this tone is itself affected by downward influences from conscious events.

A few of many examples are described below.

Mood regulation

Happiness. The “runner’s high” has been described by athletes as “pure happiness, elation, a feeling of unity with oneself and/or nature, endless peacefulness, inner harmony, boundless energy, and reduction in pain sensation” (Dietrich and McDaniel, 2004). These sensations have previously been attributed to release of endorphins (endogenous opioids) by exercise, but the evidence is equivocal since it is based entirely on measurement of endorphins in circulating blood and endorphins do not cross the blood-brain barrier. Also, endorphins depress respiration and cause pinpoint pupils, changes not seen in runners.

However, exercise (running and cycling for one hour at moderate intensity) causes a significant rise in blood levels of anandamide, which does cross the blood-brain barrier, compared with non-exercising controls (Sparling et al., 2003). Exercise makes ‘demands’ on many systems and presumably causes widespread, but specific, peripheral and central release of anandamide. Endurance athletes, such as marathon runners, also report euphoria, sedation (post-exercise calm), decreased anxiety and sense of well-being, similar to the “runner’s high” and to the effects of CB₁ receptor agonists such as THC (Sparling et al., 2003; Dietrich and McDaniel, 2004). These authors suggest that endocannabinoids, rather than endorphins, are the key mediators of the “runner’s high”. The very name anandamide is derived from the Sanskrit word *ananda* meaning bliss.

Yet both cannabinoid and opioid receptors are present in the major reward pathways in the brain – the dopaminergic mesolimbic pathway which includes the nucleus accumbens. Both CB₁ receptor agonists such as anandamide and opioids increase dopamine release in the nucleus accumbens, and this action is believed to be the basis of all rewarding and addictive drugs and of natural rewards such as eating and sex (Gardner, 1999). There is a close interaction between cannabinoids and opioids. For example, the reinforcing properties of opioids are greatly reduced in CB₁ knockout mice (Ledent et al., 1999). Both CB₁ agonists such as THC and opioids such as morphine and heroin are drugs of addiction. Their combined actions, both centrally and peripherally, may not only account for the “runners’ high” but also for the addictive quality of regular exercise and the withdrawal symptoms experienced when the exercise is stopped. A similar cannabinoid/opioid interaction is involved in analgesia (see below).

In relation to consciousness, athletes do not initially set out in order to be happy and pain-free, and the “high” does not always occur. But when it does, it seems to bubble up spontaneously into consciousness from the undercurrents of increased endocannabinoid (and probably endorphinoid) tone. It is interesting that animals appear to be primed to enjoy activities that they need for survival, such as chasing quarry, eating and sex, and man is no exception.

Depression, anxiety, psychosis. Unpleasant emotions obtrude into consciousness if the endocannabinoid system is overstimulated. For example, THC, the CB₁ receptor agonist in cannabis, has biphasic actions: low doses decrease anxiety, tension and depression, while high doses provoke anxiety, panic, dysphoria and acute psychosis. Undesirable reactions are more common in naïve subjects, stressful situations, mental illness, and chronic use (Ashton et al., 2005). These reactions are associated with smoking, especially of potent cannabis preparations such as “skunk”, or intravenous administration which flood the endocannabinoid system, in contrast to anandamide which, as mentioned above, is physiologically released according to tissue ‘demand’ and quickly inactivated.

The effects of CB₁ overstimulation have clinical implications. There is evidence from *post-mortem* studies that elevated levels of endocannabinoids (anandamide and 2-AG) and of CB₁ receptors are present in the dorsolateral prefrontal cortex of depressed suicide victims and also in alcoholic suicides compared with controls. Similar *post-mortem* changes are found in the prefrontal cortex and anterior cingulate gyrus in schizophrenics (Tzavara and Witkin, 2008). All these conditions, as well as depressed patients with Parkinson’s disease, were associated with polymorphisms of the CNRI gene which codes for CB₁ receptors (Vinod and Hungund, 2006). It is not clear whether these changes are causative or whether they indicate compensatory adaptations to initial low levels of endocannabinoids but the authors “hypothesise that the sensitisation of CB₁-receptor-mediated G protein signalling in the prefrontal cortex is one of the etiological or neuroadaptive factors in the pathophysiology of suicide”. Further discussion and potential cannabinoid therapies for depression are discussed by Tzavara and Witkin (2008).

However, this hypothesis is counterintuitive as an explanation for depression in general. Decreased CB₁ activity seems more likely to account for the anhedonia, anxiety, decreased pain tolerance, chronic pain conditions, and decreased serotonergic activity that often accompanies depression, as well as the observation that the CB₁ receptor antagonist, rimonabant, causes depression and anxiety in a significant proportion of psychiatrically normal subjects (Pi-Sunyer et al., 2006). It may be the location and degree of endocannabinoid tone that determines a person’s conscious mood.

With regard to anxiety, there is convincing preclinical evidence that stressful situations provoking anxiety increase cannabinoid tone, as shown by raised anandamide levels in the amygdala, and endocannabinoids, in turn, decrease anxiety. Endocannabinoids also exert an amnesic effect, also probably mediated through the amygdala and hippocampus, and are crucial for the extinction (forgetting) of aversive memories (Degroot, 2008). These actions are consistent with the observation that cannabis smokers often forget what they started to say before they come to the end of a sentence. Thus endocannabinoids, “the brain’s

own marijuana”, act as a kind of dimmer switch, decreasing anxiety during stress and the memory of the fear afterwards, thus helping the individual to cope with stressful situations and their after effects (perhaps even preventing post-traumatic stress disorder (PTSD)).

Observations of the effects of cannabidiol (CBD), a cannabinoid present in cannabis, has clinical implications in anxiety disorders. CBD has anxiolytic, hypnotic, antipsychotic and anticonvulsant actions, and has similar anxiolytic effects to diazepam in a simulated public speaking task (references in Ashton et al., 2005; Degroot, 2008). Its effects appear to be due to inhibition of both inactivation and reuptake of anandamide, thereby increasing anandamide concentrations. CBD may also act as an agonist at serotonin (5-HT_{1A}) receptors, similar to the effects of the SSRIs (selective serotonin reuptake inhibitors) the most widely prescribed treatment for anxiety disorders (Gobbi et al., 2005). SSRIs are not without adverse effects; drugs that modestly increase endocannabinoid tone may offer a more “physiological” approach for the treatment of both depression and anxiety disorders.

Pain modulation

Pain, of course, is only pain when it becomes conscious. But the systems underlying the modulation of pain perception remain intact and active below consciousness, particularly the powerful pain inhibitory mechanisms, many of which involve endocannabinoids. For example, potentially painful stimuli may pass unnoticed during the excitement of sporting activity (see “runner’s high” above). There is a film well known to pain researchers, in which a racehorse is seen to break its foreleg during a race. The horse gallops on regardless and wins the race. Only after passing the winning post does the valiant steed begin to limp. All sentient animals possess this innate self-analgesic mechanism which clearly has survival value when fighting or fleeing from danger.

An elaborate endocannabinoid system modulates both acute and chronic pain responses at all levels of the nervous system from the periphery to the brain (Figure 2). The system involves CB₁, CB₂ and other cannabinoid receptors, and interacts with endogenous opioid systems and descending pain suppressant pathways mediated by monoaminergic and opioid neurotransmitters.

CB₁ receptors. CB₁ receptors are present in many pain-processing pathways in the brain, including sensory cortex, thalamus, amygdala and periaqueductal grey. In rodents microinjection of CB₁ agonists into these areas suppresses pain responses, and electrical stimulation of the periaqueductal grey causes a marked increase in

release of anandamides. CB₁ receptors are also present in pain pathways in the dorsal horn of the spinal cord and systemically administered CB₁ agonists inhibit activity in spinal dorsal horn neurones. In brain and spinal cord CB₁ receptor activation decreases neuronal excitability and release of neurotransmitters (Figure 1), including pain stimulating peptides (e.g. substance P) and excitatory transmitters such as glutamate, thereby acting as for anxiety, as a dimmer switch for pain.

CB₂ receptors. CB₁ receptors are also present at peripheral terminals of primary sensory afferent nerves where they interact with CB₂ receptors and with a more recently characterised ionotropic vanilloid receptor (TRVP₁) which provokes pain when activated by capsaicin (Nagy et al., 2008). CB₂ receptors, situated in cells and tissues associated with inflammatory reactions, appear to modulate pain initiation at sites of tissue injury, producing anti-inflammatory effects by decreasing the release of nociceptive substances such as histamine, substance P and others which activate both CB₁ and CB₂ receptors.

Vanilloid receptors. TRVP₁ receptors are present at the peripheral terminals of sensory nerves and also in the dorsal root of the spinal cord, the trigeminal ganglion, the gut and elsewhere. They are activated by capsaicin and have the unique feature of responding to noxious increases in temperature. They are also activated by anandamide, a finding which seems contradictory to its analgesic effect. It is suggested that the TRVP₁/anandamide interaction represents an adaptive device which dampens the pain of moderate temperature rises caused by inflammation (through cannabinoid receptors) but increases pain (via TRVP₁ receptors) when temperatures reach levels that threaten tissue damage, at this stage facilitating a withdrawal response in intact animals (Zygmunt et al., 2000).

Cannabinoid/opioid interactions. The endocannabinoid system interacts closely with an opioid pain modulating system, producing a complex of parallel, interlaced, and complementary analgesic mechanisms. Both opioid and cannabinoid receptors are coupled to G_{i/o} proteins and decrease cAMP production (Figure 1), and the two classes of neuromodulators have additive or synergistic analgesic effects. Endocannabinoids may increase the synthesis and/or release of endogenous opioids in brain and spinal cord and some of their effects may be mediated via opioid receptors. In animals, the endocannabinoid system appears to be tonically active while the endogenous opioid system is activated by intense or prolonged stress or pain and high threshold electrical stimulation, although both systems became more active as pain intensity increases (Walker et al., 1999). The systems appear to act through different pathways since opioid antagonists do not completely block the analgesic effects of cannabinoids. Much evidence has been obtained in rodents, but a clinical example was provided by Hamann and diVadi

(1999) who described a patient with neuropathic pain due to multiple sclerosis. The pain was controlled by the synthetic cannabinoid agonist nabilone but not by opioids, and the analgesic effect of nabilone was not reversed by the opioid antagonist naloxone.

Thus endocannabinoids regulate a tonically active, vital and subconscious system that protects an animal from being overwhelmed by potentially painful stimuli. As in mood regulation, endocannabinoids interact with endogenous opioid systems. Pain itself is a protective sensation, but normally erupts into consciousness only when noxious stimuli are intense and/or cognitive decisions are required, although fear and pathological conditions, such as depression and anxiety, can overcome the “pain barrier”.

Synthesis

Only mood and pain modulation are described here, but the endocannabinoid system performs similar functions in sleep, reward, cognition and many other activities mentioned above. In fact, every conscious thought or feeling, everything we perceive, think or do (along with the relevant physical accompaniments such as adjustments in heart rate and respiration) is primed and tempered by the level of endocannabinoid tone. It is one of several systems that protect the conscious brain from being overwhelmed by too much information – more than it can attend to at any one moment. In subtle ways endocannabinoids help to form our personality – whether we are clear-thinking and focussed, laid back and relaxed or over-attentive, anxious and fearful. Without the system we would be mentally confused, less joyful, more anxious, sleepless, and more affected by pain.

In this capacity the endocannabinoid system interacts with several other modulator systems which also seethe in the undercurrent depths. The brain is richly endowed with multiple back-up circuits, each performing similar functions utilising different but mutually reinforcing anatomical and chemical pathways. As described, pain can be suppressed through discrete noradrenergic, serotonergic, dopaminergic, opioid or endocannabinoid pathways (Figure 2). The particular pain suppressant system or combination of systems activated may depend partly on the type of stress to which the animal is exposed. This elaborate organisation of copious back-up systems appears to be a feature of many, if not all, vital life-sustaining processes (Ashton, 1992).

But it is a two way process. Influences from unconscious depths may provide a “gut feeling” about a situation; those that bubble up into consciousness may be suppressed or reinforced by downward neural activity from the conscious brain. It is possible to suppress unwanted thoughts by conscious effort (“Shut up, I don’t

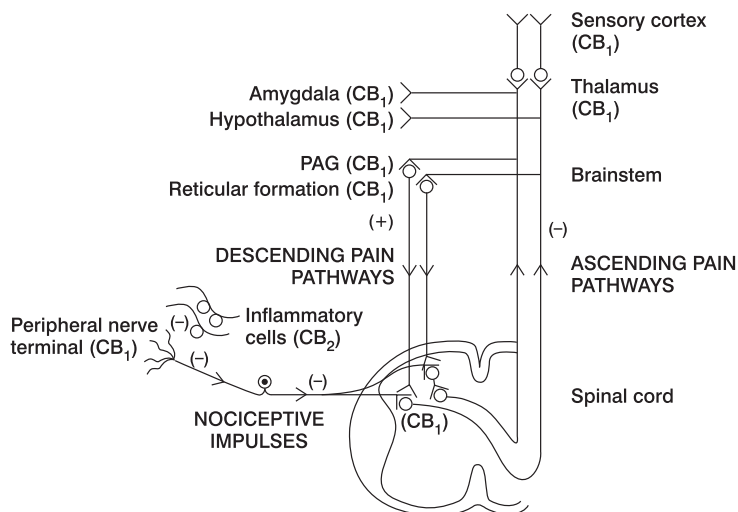


Figure 2. Cannabinoid receptors in pain processing pathways.

Cannabinoid CB₁ receptors are present on peripheral nerve terminals, dorsal horn spinal cord, and pain processing pathways in the brain. Their activation inhibits activity in excitatory ascending pain pathways and stimulates activity in descending inhibitory pain pathways. CB₂ receptors are present in blood cells associated with inflammation and may modulate pain initiation at site of tissue injury. (PAG = periaqueductal grey matter).

want to think about that”) or to encourage them (“Yes, come on, this may require a big decision”). Reciprocal impulses pass all the time between conscious and below conscious neural networks. The part played by the endocannabinoid system at all levels of consciousness has up to now been largely ignored. But, it is clear that the “mind” stems not just from consciousness but also critically depends on the undercurrent of non-conscious processes.

References

- Alger, B. E. (2004). Endocannabinoids: Getting the message across. *Proceedings of the National Academy of Sciences of the USA*, 101, 8512–8513.
- Ashton, C. H., Moore, P. B., Gallagher, P. and Young, A. H. (2005). Cannabinoids in bipolar affective disorder: A review and discussion of their therapeutic potential. *Journal Psychopharmacology*, 19, 293–300.
- Ashton, H. (1992). Overview. In *Brain Function and Psychotropic Drugs* (pp. 3–4). Oxford: Oxford University Press.

- De Petrocellis, L., Cascio, M. G. and DiMarzo, V. (2009). The endocannabinoid system: A general view and latest additions. *British Journal of Pharmacology*, 141, 765–774.
- Degroot, A. (2008). Role of cannabinoid receptors in anxiety disorders. In A. Kofalvi (Ed.), *Cannabinoids and the Brain* (pp. 559–72). New York, NY: Springer Science.
- Dietrich, A. and McDaniel, W. F. (2004). Endocannabinoids and exercise. *British Journal of Sports Medicine*, 38, 536–541.
- Gardner, E. L. (1999). Cannabinoid interaction with reward systems. In G. G. Nahas, K. M. Sutin, D. J. Harvey and S. Agurell (Eds.), *Marihuana and Medicine* (pp. 187–205). Totowa, New Jersey: Humana Press.
- Gobbi, G., Bambico, F. R., Mangieri, R., Bortolato, M., Campolongo, P., Solinas, M., Cassano, T., Morgese, M. G., Debonnel, G., Duranti, A., Tontini, A., Tarzia, G., Mor, M., Trezza, V., Goldberg, S. R. and Cuomo, V. (2005). Antidepressant-like activity and modulation of brain monoaminergic transmission by blockade of anandamide hydrolysis. *Proceedings of the National Academy of Sciences of the USA*, 102, 18620–18625.
- Hamann, W. and diVadi, P. P. (1999). Analgesic effect of the cannabinoid analogue nabilone is not mediated by opioid receptors. *Lancet*, 353, 560.
- Kofalvi, A. (2008). Editorial. In A. Kofalvi (Ed.), *Cannabinoids and the Brain* (pp. v–viii). New York, NY: Springer Science.
- Ledent, C., Valverde, O., Cossu, G., Petitot, F., Aubert, J-F, Beslot, F., Bohme, G. A., Imperato, A., Pedrazzini, T., Roques, B. P., Vassart, G., Fratta, W. and Parmentier, M. (1999). Unresponsiveness to cannabinoids and reduced addictive effects of opiates in CB₁ receptor knock-out mice. *Science*, 283, 401–404.
- Nagy, I., White, J. P. M., Paule, C. C., Maze, M. and Urban, L. (2008). Functional molecular biology of the TRPV₁ ion channel. In A. Kofalvi (Ed.), *Cannabinoids and the Brain* (pp. 101–30). New York, NY: Springer Science.
- Perry, E. and Young, A. (2002). Neurotransmitter networks. In E. Perry, H. Ashton and A. Young (Eds.), *Neurochemistry of Consciousness* (pp. 3–24). Amsterdam: John Benjamins.
- Pertwee, R. G. (2001). Cannabinoid receptors and pain. *Progress in Neurobiology*, 63, 569–611.
- Pi-Sunyer, F. X., Aronne, L. J., Heshmati, H. M., Devin, J. and Rosenstock, J. (2006). Effect of Rimonabant, a cannabinoid-1 receptor blocker, on weight and cardiometabolic risk factors in overweight and obese patients. *Journal of the American Medical Association*, 295, 761–75.
- Sparling, P. B., Giuffrida, A., Piomelli, D., Rosskopf, L. and Dietrich, A. (2003). Exercise activates the endocannabinoid system. *Neuroreport*, 14, 2209–2011.
- Tzavara, E. T. and Witkin, J. M. (2008). The cannabinoid controversy: Cannabinoid agonists and antagonists as potential novel therapies for mood disorders. In A. Kofalvi (Ed.), *Cannabinoids in the Brain* (pp. 529–58). New York, NY: Springer Science.
- Vinod, K. Y. and Hungund, B. L. (2006). Role of the endocannabinoid system in depression and suicide. *Trends in Pharmacological Sciences*, 27, 539–545.
- Walker, J. M., Huang, S. M., Strangman, N. M., Tsou, K. and Sanudo-Pena, M. C. (1999). Pain modulation by release of the endogenous cannabinoid anandamide. *Proceedings of the National Academy of Sciences of the USA*, 96, 12198–12203.
- Zygmunt, P. M., Julius, D., Di Marzo, V. and Hogenstatt, E. D. (2000). Anandamide – The other side of the coin. *Trends in Pharmacological Sciences*, 21, 43–44.

Disconnecting consciousness

The neuroscience of general anaesthesia

Lucas Campos, Adrian Pichurko and George A. Mashour

We discuss the molecular and cognitive mechanisms of general anaesthesia and their relevance to the science of consciousness. General anaesthetics serve as an important tool to explore the transitions from conscious to unconscious processing, as these drugs rapidly induce a complete and reversible cessation of consciousness. With 40,000,000 patients undergoing general anaesthesia each year in North America alone, the highly controlled and monitored setting of the operating room can become the ultimate consciousness laboratory.

Introduction

Anaesthesiologists confront the possibility of intraoperative awareness millions times per year. Lack of subjective awareness is the very definition of the word “anaesthesia,” which is rooted in the Greek *aisthēsia* meaning sensation or perception. The suggestion for use of the term “anaesthesia” was made by physician-poet Oliver Wendell Holmes to describe the state of mind induced by the ether that William Morton administered as a surgical anaesthetic on October 16, 1846, at the Massachusetts General Hospital. Despite the successful clinical use of general anaesthetics for more than 150 years, the mechanism for the hypnotic component of anaesthesia is still unknown. However, the problem at the root of this mystery has been debated by philosophers since antiquity and is now the focus of intense neuroscientific investigation – we refer to the problem of consciousness.

Neuroscientific investigation is currently focused on elucidating the neural correlates of consciousness (NCC), defined by Crick and Koch as the minimal neuronal mechanisms jointly sufficient for any one specific conscious percept (see Crick and Koch, 2003). In a related line of investigation, anaesthesiologists are concerned with the neural correlates of unconsciousness. It is through these pursuits that each discipline could advance the other. In their endeavour to find the NCC, neuroscientists have traditionally examined both the state or level of consciousness (e.g., mechanisms for being unconscious or conscious) and the content

of consciousness (e.g., mechanisms for the perception of a face vs. a house). To study the state of consciousness, neuroscience has been relegated to examining patients with varying degrees of brain damage or congenital defects who are in a vegetative state. In order to study the content of consciousness, researchers can only study subjects who are already conscious. The main criticism for content-based research is: how can you find the NCC in subjects that are already conscious? Aren't these the neural correlates of perception or awareness rather than consciousness (see Hohwy, 2009)? The main criticism for state-based investigators is: Don't you risk being overly inclusive in your identification of NCC if you are simply looking at neural patterns of activity in patients that do not regain consciousness? How do you know which patterns to ignore and which are truly NCC (Hohwy, 2009)?

Anaesthesiologists regularly initiate a deep state of *reversible* unconsciousness with a variety of pharmacologically distinct drugs, which provides a unique opportunity for understanding the mechanisms of consciousness. Indeed, the operating room could be considered the ultimate consciousness laboratory. Reciprocally, discovery of the NCC will help advance our understanding of anaesthetic substrates, facilitating the generation of more specific and targeted anaesthetics, as well as better monitoring to ensure adequate depth of anaesthesia during surgery. In our examination of how these two fields can facilitate each other we will review what is known about the neural correlates of general anaesthesia, including the molecular, neuroanatomic and neurophysiologic targets, as well as how the NCC fit into this understanding. Although the idea that the study of consciousness and anaesthesia share a common goal is not new (see Hameroff, 2006; John and Pritchep, 2005; Mashour, 2004; Mashour, 2006), the thinking, understanding, and resources are finally in place for some truly exciting research to begin.

Molecular targets

Molecular targets of anaesthesia include synaptic receptors for neurotransmitters such as γ -Aminobutyric acid (GABA), glutamate and acetylcholine, as well as potassium leak channels such as TASK and TREK. These receptors are crucial to the proper functioning of numerous neural circuits. Many anaesthetics act by increasing GABA receptor activity, which hyperpolarizes neurons and makes them less responsive to propagating neural signals (see Franks, 2008). However, the mechanism of general anaesthetics is not that specific – many have multiple receptor-mediated effects, which results in depression of the cardiovascular and respiratory systems as well. We will focus on GABA as this is the most consistently activated receptor with major potentiation by barbiturates, propofol, etomidate,

isoflurane, sevoflurane, and desflurane, as well as minor potentiation by nitrous oxide and ketamine (see Alkire et al., 2008).

GABA receptors are divided into GABA-A (Ligand-gated ion channel) and GABA-B (G-protein coupled receptor) each with their own subtypes. GABA-A is the target of the majority of anaesthetic drugs and consists of five of the 19 known subunits that can be matched in various combinations to form the receptor complex. Changing one subunit affects which anaesthetic will bind and how strongly this association will be. For example, GABA-A with a beta-3 subunit mutation greatly diminishes the effectiveness of propofol and etomidate in producing unconsciousness in mice (see Orser, 2007).

An exciting finding is the identification of extrasynaptic GABA-A receptors on the neuronal cell body and dendrites that contain the alpha-5 subunit and are so sensitive they can be activated by the tiny amount of GABA neurotransmitter present in the extracellular space after activation from a nearby synapse (Orser, 2007). When activated, these receptors cause a low level of diffuse inhibitory current all along the cell body, as opposed to focally at the synapse. During *in vitro* experiments, brain slices exposed to concentrations of propofol, etomidate, and isoflurane several times lower than would be required to block transmission at the synapse amplified this low-level inhibitory current by as much as 35-fold, hyperpolarizing the cells (Orser, 2007). These novel receptors are found in high concentrations in the thalamus, hippocampus and cortical areas. The alpha-5 subunit of these extrasynaptic receptors can be exploited for drug development. Such compounds would obviate the need for high concentrations of current anaesthetic drugs that suppress the respiratory and cardiovascular systems as a side-effect to the suppression of consciousness (Orser, 2007).

Neuroanatomic targets

Knowledge of potential molecular targets for general anaesthesia can be refined using knowledge of which neuroanatomic areas are involved in consciousness. Is a target in the thalamus better than one in the cerebellum? Neuroscientific research reveals some important neuroanatomic targets known to be inhibited by anaesthesia including sleep-wake centers in the brainstem, thalamus, lateral frontoparietal system, and precuneus (Mashour, 2006).

Loss of consciousness induced by anaesthesia consistently involves a reduction of thalamic metabolism and blood flow. The thalamus relays incoming sensory information from the periphery to the cerebral cortex, passing through the thin overlying layer known as the thalamic reticular nucleus. Interestingly, the thalamus receives most of its inputs from the cortex through a massive network

of corticothalamic feedback fibres. Injecting GABA into intralaminar nuclei inhibits these fibres and rats rapidly lose consciousness, while rats anesthetized with sevoflurane can be awakened by injecting nicotine into the centromedian thalamus (see Alkire et al., 2007). Consistent with this potential mechanism for anaesthesia, deep brain electrical stimulation of the central thalamus improved behavioral responsiveness in a patient who was in a minimally conscious state (see Schiff et al., 2007).

The thalamus, however, is not a simple switch for consciousness as previously thought. Superficial scalp and deep subthalamic electrophysiological recordings during the induction of anaesthesia showed cortical electrical activity decreases in association with loss of consciousness, while decreases in thalamic activity were delayed and more associated with immobility (see Velly et al., 2007). In experiments in which the thalamus was removed, the cortex continued to produce an activated EEG, while removal of the cortex prevented anaesthetics from decreasing the metabolic and electrophysiological activity in the thalamus (see Nakakimura et al., 1988). Collectively, these experiments suggest that the thalamus reflects global cortical activity in addition to modulating its function.

Neuroimaging demonstrates deactivation of the precuneus (medial parietal lobe), mesial parietal and posterior cingulate cortex in anaesthetic-induced unconsciousness as well as in vegetative patients (see Laureys et al., 2006). Using intermediate doses of nitrous oxide causes feelings of depersonalization and out-of-body experiences in association with deactivation of the posterior mesial cortex. Activity in the lateral temporo-parieto-occipital complex centred on the inferior parietal cortex is also altered with anaesthesia (Alkire et al., 2008). Therefore, activation of unique molecular targets in these specific brain areas may suppress consciousness. We can, however, be more specific and ask: What type of neural physiology has to be inhibited for anesthetic-mediated loss of consciousness to take place?

Functional neurophysiologic targets

Among the major hypotheses investigated regarding the NCC are synchronous activity and re-entrant or feedback activity. The synchrony hypothesis says neurons have to coordinate their output, typically in the gamma range (30–70 Hz), in order for it to be consciously represented. If you are looking at a blue circle, for example, the neurons recognizing a circle and those recognizing blue have to synchronize their outputs so that the two are bound together into the single percept of a blue circle. The role of gamma activity in anesthetic-mediated unconsciousness is unclear, as gamma activity can actually be augmented during evoked potentials under general anaesthesia (see Imas et al., 2004). Thus, it is not simply

the case that gamma activity is suppressed during general anaesthesia, but likely that the capacity for such waves to meaningfully integrate neural information is disrupted. In support of this, mean integration capacity was most prominently reduced in gamma band after propofol induction and this breakdown of spatio-temporal organization of gamma waves (reflecting loss of integration) correlated with loss of consciousness (see Lee, Mashour et al., 2009).

“Re-entrant” neuronal activity refers to a reverberant wave of activity that follows an initial impulse of neural activity. When you see a face, for example, the visual neurons release a volley of feedforward sensory output to the cortex, but the face will not be recognized until a feedback volley of modulatory output from a higher area such as the frontal cortex is detected by the fusiform facial area. The importance of this kind of re-entrant activity was demonstrated in an experiment in which human subjects were given propofol and feedforward and feedback neural activity was assessed by EEG (see Lee, Kim et al., 2009). It was found that when subjects were conscious, the feedback activity was significantly higher than the feedforward activity. When the subjects lost consciousness, the feedback activity decreased significantly while the feed forward activity rebounded quickly, suggesting that sensory information was still being received by the brain despite the loss of consciousness (Figure 1) (Lee, Kim et al., 2009). This supports earlier work showing feedback inhibition during anaesthesia in an animal model (see Imas et al., 2005).

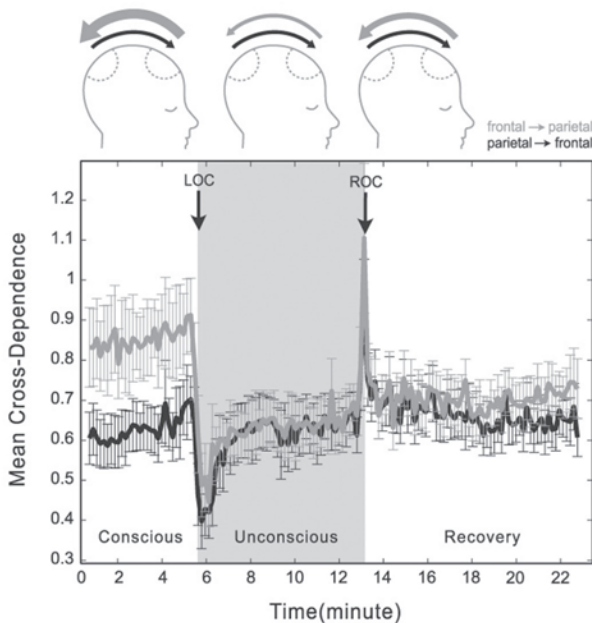


Figure 1

What is the mechanism of this inhibition of feedback circuitry? In the re-entrant example, we know that propofol targets GABA-A and that there is a higher density of the exquisitely sensitive extra-synaptic GABA-A receptors in the thalamus (which acts a hub for cortical input). Could activation of extra-synaptic GABA-A receptors cause a desynchronization in the gamma range that disrupts the feedback output or diminishes the amount of information being integrated? Experiments using high doses of several anaesthetics found these drug levels produced a stereotypical burst-suppression pattern in which the EEG went from flat to global activation at regular intervals. Thus, the corticothalamic system lost its usual repertoire of responses to incoming information and was left with being either on or off. These findings from studies of general anaesthesia suggest that the level of consciousness is directly related to the ability to integrate the massive amount of neural information sent to the brain (Alkire et al., 2008).

A measure of integrated information has been proposed called phi (Φ) that quantifies the integration of information. The more information a system is able to integrate the higher the value of phi. The information integration theory of consciousness explains why the corticothalamic system is important to modulate in order to produce unconsciousness: these concentrated connections of specialized neurons are involved in a tremendous amount of information integration from nearly all regions of the brain. The parallel theories in anaesthesiology suggest that it is the unbinding or lack of integration that is associated with anesthetic-induced unconsciousness (Mashour, 2004; 2006). We are in the early phase of proving these theories, but more and more data are supportive (see Imas et al., 2005; Lee, Kim et al., 2009; Lee, Mashour et al., 2009; Peltier et al., 2005). The information integration theory of consciousness and its corollary in anaesthesiology has implications for the boundary of conscious and nonconscious processes: instead of a binary framework, this theory suggests gradations of conscious activity. Although the induction of general anaesthesia is typically accomplished rapidly in the operating room, the study of different steady state levels at varying degrees of sedation may help elucidate the neurobiology of this transition.

Conclusion

We have examined the molecular, anatomic and neurophysiologic targets of anaesthesia and their role in the mechanism of general anaesthesia. Currently, the non-specific effects of general anaesthetics are associated with consequences for the cardiovascular and respiratory systems. Such physiologic depression often leads the anaesthesia provider to decrease anaesthesia levels, which creates the

possibility of intraoperative awareness. This suggests the need for improved, rationally designed anaesthetics that target the NCC more specifically.

Anesthesiology is in a unique position to contribute to understanding consciousness because it is a controlled, reversible, and intentional suppression of the NCC at some level. With a combined endeavour, neuroscience and anaesthesiology will be able to pioneer new anesthetic drugs and monitors of consciousness as we evolve our understanding from mere correlation to causation of conscious states. This discovery will help complete the revolution created by Morton's first demonstration of general anaesthesia in 1846 – with the use of general anaesthetics we will hopefully elucidate how the biological machinery of the brain produces the subjective experiences of the mind.

References

- Alkire, M. T., Hudetz, A. G. and Tononi, G. (2008). Consciousness and anesthesia. *Science*, 322, 876–880.
- Alkire, M. T., McReynolds, J. R., Hahn, E. L. and Trivedi, A. N. (2007). Thalamic microinjection of nicotine reverses sevoflurane-induced loss of righting reflex in the rat. *Anesthesiology*, 107, 264–272.
- Crick, F. and Koch, C. (2003). A framework for consciousness. *Nature Neuroscience*, 6, 119–126.
- Franks, N. P. (2008). General anaesthesia: From molecular targets to neuronal pathways of sleep and arousal. *Nature Reviews Neuroscience*, 9, 370–386.
- Hameroff, S. R. (2006). The entwined mysteries of anesthesia and consciousness: Is there a common underlying mechanism? *Anesthesiology*, 105, 400–412.
- Hohwy, J. (2009). The neural correlates of consciousness: New experimental approaches needed? *Consciousness and Cognition*, 10, 101–106.
- Imas, O. A., Ropella, K. M., Ward, B. D., Wood, J. D. and Hudetz, A. G. (2005). Volatile anaesthetics disrupt frontal-posterior recurrent information transfer at gamma frequencies in rat. *Neuroscience Letters*, 387, 145–150.
- Imas, O. A., Ropella, K. M., Wood, J. D. and Hudetz, A. G. (2004). Halothane augments event-related gamma oscillations in rat visual cortex. *Neuroscience*, 123, 269–278.
- John, E. R. and Pritchep, L. S. (2005). The anesthetic cascade: A theory of how anesthesia suppresses consciousness. *Anesthesiology*, 102, 447–771.
- Laureys, S., Boly, M. and Maquet, P. (2006). Tracking the recovery of consciousness from coma. *Journal of Clinical Investigation*, 116, 1823–1825.
- Lee, U., Kim, S., Noh, G. J., Choi, B. M., Hwang, E. and Mashour, G. A. (2009). The directionality and functional organization of frontoparietal connectivity during consciousness and anesthesia in humans. *Consciousness and Cognition* 18, 1069–1078.
- Lee, U., Mashour, G. A., Kim, S., Noh, G. J. and Choi, B. M. (2009). Propofol induction reduces the capacity for neural information integration: implications for the mechanism of consciousness and general anesthesia. *Consciousness and Cognition*, 18, 56–64.

- Mashour, G. A. (2004). Consciousness unbound: Toward a paradigm of general anesthesia. *Anesthesiology*, 100, 428–433.
- Mashour, G. A. (2006). Integrating the science of consciousness and anesthesia. *Anesthesia and Analgesia*, 103, 975–982.
- Nakakimura, K., Sakabe, T., Funatsu, N., Maekawa, T. and Takeshita, H. (1988). Metabolic activation of intercortical and corticothalamic pathways during enflurane anesthesia in rats. *Anesthesiology*, 68, 777–782.
- Orser, B. A. (2007). Lifting the fog around anesthesia. *Scientific American*, June, 54–61.
- Peltier, S. J., Kerssens, C., Hamann, S. B., Sebel, P. S., Byas-Smith, M. and Hu, X. (2005). Functional connectivity changes with concentration of sevoflurane anesthesia. *Neuroreport*, 28, 285–288.
- Schiff, N. D., Giacino, J. T., Kalmar, K., Victor, J. D., Baker, K., Gerber, M., Fritz, B., Eisenberg, B., Biondi, T., O'Connor, J., Kobylarz, E. J., Farris, S., Machado, A., McCagg, C., Plum, F., Fins, J. J. and Reza, A. R. (2007). Behavioral improvements with thalamic stimulation after severe traumatic brain injury. *Nature*, 448, 600–603.
- Velly, L. J., Rey, M. F., Bruder, N. J., Gouvitsos, F. A., Witjas, T., Regis, J. M., Peragut, J. C. and Gouin, F. M. (2007). Differential dynamic of action on cortical and subcortical structures of anesthetic agents during induction of anesthesia. *Anesthesiology*, 107, 202–212.

Consciousness and neural time travel

Michael E. Hasselmo

Consciousness is described in this article as a perception of self that is constructed for guidance of voluntary goal-directed behaviour, including neural mechanisms for mental time travel into the future during planning and into the past during episodic retrieval. Neural time travel draws on the distributed cortical perceptual processes for detecting the state of the self along multiple dimensions, including spatial location, head direction, speed, temporal duration, and egocentric relationships to items. These circuit mechanisms are regulated by neuromodulatory influences such as muscarinic acetylcholine receptors that activate intrinsic properties of neurons involved in maintaining an internal representation of self.

In this article, I will argue that consciousness is the perception of self that is constructed as a necessary component of an internal model for understanding and selecting one's own actions, including the planning and generation of voluntary movement in goal-directed behavior, as well as the retrieval of episodic memories of prior behavior. This perception of self is required for understanding current state, as well as projecting the self into future situations, or retrieving memories of the self in past episodes. I will refer to these processes as neural time travel, providing the biological mechanisms for the mental time travel originally defined as part of episodic memory (Tulving, 2002). This perception of a conscious self is a distributed process mediated by cortical structures which provide conscious perception along a variable number of continuous dimensions, and are diffusely regulated by neuromodulatory systems of the brain.

This definition arises from my work on modelling of the neural mechanisms of episodic memory and the planning and generation of future goal-directed behavior. My models utilize an internal representation of the location of self in space and time (Hasselmo 2008b; Hasselmo, 2009), both during the retrieval of segments of spatiotemporal trajectories from the past, and during planning of future or imaginary behavior. These processes require a specific neural representation of the state of self that has been described for physical quantities such as location, heading and speed (Hasselmo 2008b; Hasselmo, 2009), but could also

include eye position, limb orientation, current goals, attentional focus on specific sensory information and emotional or drive state. This internal representation of self state can be updated by current chosen actions and by external sensory input, or in an introspective representation of self, can be updated by internally generated planned actions, or by selection of remembered cues and actions during retrieval of episodic memory. Thus, the updating of self state forms a component of mechanisms for neural time travel to re-experience past states or evaluate possible future states.

Perceiving the self in neural time travel

This definition of consciousness arises from analysis of potential neural mechanisms for encoding and retrieval of episodic memory, and planning of current or future actions. These processes require the capacity for altering the state of an internal model of self, by altering position on the time dimension (for mental time travel into past events or possible future events), or by generating mental actions that evaluate alterations in other dimensions of self-state. For example, while writing this article, I must utilize an internal representation of my self as an agent evaluating these processes. I have focused on specific neural mechanisms, but regardless of the neural mechanisms, the evaluation of these processes in computer simulations requires a computational representation of an agent's self-state. For example, in simulating episodic memory in an agent, I simulated the agent's head direction, speed, allocentric location and relative distance from items during behavior and then these variables were recapitulated during episodic retrieval (mental time travel into the past). Just as I found a representation of an agent's self-state necessary to simulate processes of episodic memory and planning, conscious individuals utilize an internal representation of self-state as part of performing these same processes.

Note that I am using terms like "state" and "agent" based on reinforcement learning theory (Sutton and Barto, 1998). The framework of reinforcement learning theory has proven effective as a mathematical framework for modelling operant behavior, both for understanding animal behavior and for guiding the movement of robots. This framework focuses on the behavior of an agent interacting with an environment. The agent chooses actions based on its current sensory state. Some choices can be made just on the basis of current input. Other behavioural choices cannot be made on the basis of current sensory input alone, but require disambiguation by memory of prior states or actions (Zilli and Hasselmo, 2008). Similarly, consciousness serves a role for other choices that require the agent to have an internal representation of self. It is an interesting question whether this

self state is innately wired in the brain, or whether it requires self-organization of neural connectivity in response to early experience. An individual brain exposed to the sensory input received by multiple different people in early development might develop an internal representation of more than one self.

Potential neural mechanisms of consciousness

As an agent describing simulations above, I noted that a representation of the self-state is necessary regardless of the nature of neural representation. However, there are specific neural mechanisms that are potential candidates for maintenance of self-state elements such as location and direction. These include rate coding representations of self-state that consist of persistent neural spiking activity (Hasselmo and Stern, 2006; Hasselmo 2008a), based on data for intrinsic persistent spiking mechanisms in individual neurons (Egorov et al., 2002; Yoshida et al., 2008). An alternative way to represent self-state is the current state of neural spiking relative to the phase of network oscillations (Hasselmo, 2008; Hasselmo et al., 2009). based on data for intrinsic mechanisms of oscillations in entorhinal cortex (Alonso and Llinas, 1989; Giocomo et al., 2007). This representation of current state by relative phase of spiking to network oscillations is based on models of theta phase precession in the hippocampus (O'Keefe and Recce, 1993) and models of grid cell firing in the entorhinal cortex (Burgess et al., 2007).

Both these cellular processes provide mechanisms for maintaining a current self state, and also for updating this self-state on the basis of neural activity induced by sensory input, and also by neural activity due to internal generation of mental actions. The generation of mental actions may sound like it is attributing a separate entity for the generation of those actions. However, internally generated mental actions can be generated by deterministic or random elements of the current self-state. For example, my self-state generates mental actions planning this sentence within a previously constructed mental landscape created by prior physical and mental activity. The next phrase is generated by components of my self-state maintaining memory of the next concept in my previously planned sequence of concepts held in my working memory. If the rate or phase code of this working memory maintenance is lost due to a distraction, the absence of a next mental action generates a search of long-term memory for earlier written notes. These earlier written notes arose from previous sensory input that inspired a novel shift in mental self-state during some earlier period of thinking.

In summary, physical and mental actions require a representation of self to code the context of observations in terms of variables such as allocentric location, head direction, time, emotional state, evaluation of current status and future

planning. Even in abstract thought, we need a representation of self as an agent generating internal mental actions such as encoding and retrieving, or even thinking and planning.

It is perhaps not an accident that the easiest example of this process utilizes Cartesian coordinates to represent the allocentric location of the self as an agent in an environment, and to represent the location of items perceived by visual or somatosensory input. In this context, the phrase 'I think therefore I am,' can be rephrased as: I construct a representation of self-state in order to represent and understand the act of thinking.

Influences on neural mechanisms

The above description of neural mechanisms for consciousness provide a perspective on physical influences on consciousness. Muscarinic cholinergic agonists can regulate a rate code for self state by activating intrinsic mechanisms of persistent spiking in neurons of the entorhinal cortex (Egorov et al., 2002; Yoshida et al., 2008) as well as in the prefrontal cortex (Haj-Dahmane and Andrade, 1999) and in the postsubiculum (Yoshida and Hasselmo, 2009). Blockade of muscarinic cholinergic receptors by antagonists such as scopolamine or atropine will reduce or prevent this persistent spiking activity, thereby impairing the internal representation of self-state. This serves to impair the encoding of self-state into episodic memory (Hasselmo and Stern, 2006) as supported by extensive literature on the impact of muscarinic cholinergic antagonists on encoding of memory (Ghoneim and Mewaldt, 1977; Hasselmo and McGaughy, 2004).

This does not mean that consciousness only involves memory. Blockade of muscarinic receptors will also impair the representation of attention to sensory stimuli maintained in the self-state, impairing attention to stimuli (Hasselmo and McGaughy, 2004) and at higher doses causing delirium (Perry and Perry, 1995). It is striking that muscarinic cholinergic antagonists are capable of completely blocking coherent thought. Blockade of other modulatory systems such as the noradrenergic, serotonergic and dopaminergic systems do not cause such a complete blockade of cognitive processes. Notably, acetylcholine levels go quite low during slow wave sleep, but rise again during the narrative dreaming of REM sleep, when conscious self-awareness is present (and yet lost to the long-term representation of self due to a difference in neural mechanisms resulting in reduced encoding).

Other drugs that cause a complete blockade of cognitive processes include the GABAergic agonists such as the benzodiazepines, the barbiturates (such as Thiopental, which for many years was a standard agent for induction of anesthesia),

and other drugs known to induce unconsciousness at higher doses such as alcohol and chloral hydrate. The impact of these drugs could be due to enhanced inhibition suppressing the persistent spiking rate code for working memory of self-state. However, these effects might also involve the alteration of the gamma frequency oscillations dependent upon GABAergic feedback. This alteration might disturb a gamma frequency phase code of neocortical activity that disrupts the maintenance of self-state. The loss of consciousness caused by anesthetics might also involve influences on the hyperpolarization activated cation current, or H current (Chen et al., 2009). The effect on H current could reduce the phase code for representation of self-state, and thereby suppress consciousness. This is only a general sketch of possible mechanisms for the direct disruption of cortical circuitry mediating conscious awareness. There are many other complex neural influences of these drugs that could contribute to the disruption of conscious processing in the neocortex.

Distributed consciousness in the neocortex

Prior empirical results suggest that consciousness is a purely cortical process, coded in a distributed manner throughout the neocortex. Data on brain injuries indicates a role of neocortical activity in consciousness. For example, the loss of cortical visual pathways with retention of subcortical visual pathways in the superior colliculus can result in the phenomenon of non-conscious “blindsight”.

The distributed nature of consciousness is also evident from effects of brain damage, including the striking effects of neglect found with parietal lesions. This clearly illustrates the role of cortical maintenance of self-state in conscious behavior. Parietal cortex plays an important role in tracking of self-state in spatial location and plotting the update in self-state associated with specific actions. A lesion of parietal cortex wipes out specific subspaces of the Cartesian coordinates of self, quantitatively reducing the scope of conscious awareness. The person experiencing this does not sense a loss of the personal perception of a self, but loses awareness of the interaction of the self with that component of their egocentric world, in a sense reducing the scope of their self-state. This is similar to the manner in which a lesion of primary visual cortex removes conscious awareness of that section of the visual field. There is no conscious perception of a black portion of the visual field, but a loss of awareness similar to the lack of awareness of visual objects behind the head.

This evidence undermines the common intuitive sense of consciousness as a unified singular process. The ability to lose a variable component along individual dimensions of conscious awareness indicates that consciousness is a distributed

and continuous quantity. Normally, the brain enters conscious or unconscious states in a relatively synchronous manner (usually most of the cortex is simultaneously engaged in the slow waves of deep sleep or anesthesia or in the low amplitude, high frequency desynchronized activity of conscious waking or dreaming sleep), and we therefore perceive ourselves as transitioning into and out of consciousness as a whole. However, if we were to selectively anesthetize progressive portions of the cortex, we would experience our conscious selves being disassembled like a jigsaw puzzle. This is done in a simple form with the sodium amytal anesthetization of one cerebral hemisphere for testing of language lateralization (the Wada test), resulting in loss of conscious use of language function.

Speculations

Similar to the progressive loss of consciousness along individual sensory and spatial dimensions, it is possible to also consider the continuous nature of consciousness across time. The local excitatory and inhibitory circuits may only be able to maintain consciousness for one gamma cycle at a time. The rate code or phase codes described above may provide working memory that allows continuity of consciousness over hundreds of milliseconds to seconds or minutes. However, the sense of conscious self in the morning as the same entity from the previous night requires long-term memory. Without long-term memory, we are discrete temporally localized conscious entities that die after each moment, and are replaced by a different self-sustained neural self-state. This is reminiscent of the Jerry Seinfeld routine where he talks about “night-time Jerry” not caring about whether “morning Jerry” will feel bad.

On the furthest extreme of speculation, I would like to mention that the anthropic principle of physics indicates that the current physical laws of the universe take their specific form because they specifically result in the existence of human beings as observers of the universe. To extend this principle to conscious self-state would logically mean that each individual’s universe has physical characteristics that specifically result in that individual conscious self-state perceiving the universe at a given point in time (with unfortunately no guarantee for future self-state).

Acknowledgements

Research reviewed here was supported by Silvio O. Conte Center grant NIMH MH71702, NIMH R01 60013, NIMH MH61492, NSF Science of Learning Center CELEST SBE 0354378.

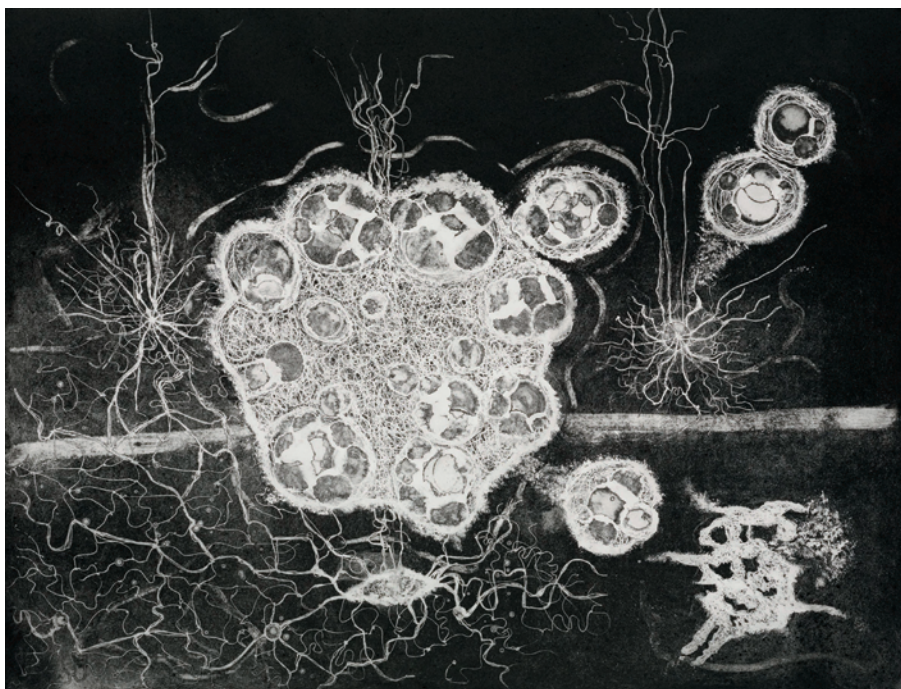
References

- Alonso, A. and Llinas, R. R. (1989). Subthreshold Na-dependent theta-like rhythmicity in stellate cells of entorhinal cortex layer II. *Nature*, 342, 175–177.
- Burgess, N., Barry, C. and O'Keefe, J. (2007). An oscillatory interference model of grid cell firing. *Hippocampus*, 17, 801–812.
- Chen, X., Shu, S., Kennedy, D.P., Willcox, S. C. and Bayliss, D. A. (2009). Subunit-specific effects of isoflurane on neuronal Ih in HCN1 knockout mice. *Journal of Neurophysiology*, 101, 129–140.
- Egorov, A. V., Hamam, B. N., Fransén, E., Hasselmo, M. E. and Alonso, A. A. (2002). Graded persistent activity in entorhinal cortex neurons. *Nature*, 420, 173–178.
- Ghoneim, M. M. and Mewaldt, S. P. (1977). Studies on human memory: The interactions of diazepam, scopolamine, and physostigmine. *Psychopharmacology*, 52, 1–6.
- Giocomo, L. M., Zilli, E. A., Fransén, E. and Hasselmo, M. E. (2007). Temporal frequency of subthreshold oscillations scales with entorhinal grid cell field spacing. *Science*, 315, 1719–1722.
- Haj-Dahmane, S. and Andrade, R. (1999). Muscarinic receptors regulate two different calcium-dependent non-selective cation currents in rat prefrontal cortex. *European Journal of Neuroscience*, 11, 1973–1980.
- Hasselmo, M. and McGaughy, J. (2004). High acetylcholine levels set circuit dynamics for attention and encoding and low acetylcholine levels set dynamics for consolidation. *Progress in Brain Research*, 145, 207–231.
- Hasselmo, M. E. (2008a). Grid cell mechanisms and function: Contributions of entorhinal persistent spiking and phase resetting. *Hippocampus*, 18, 1213–1229.
- Hasselmo, M. E. (2008b). Temporally structured replay of neural activity in a model of entorhinal cortex, hippocampus and postsubiculum. *European Journal of Neuroscience*, 28, 1301–1315.
- Hasselmo, M. E. (2009). A model of episodic memory: Mental time travel along encoded trajectories using grid cells. *Neurobiology Learning and Memory*, 92, 559–573.
- Hasselmo, M. E., Brandon, M. P., Yoshida, M., Giocomo, L. M., Heys, J. G., Fransen, E., Newman, E. L. and Zilli, E. A. (2009). A phase code for memory could arise from circuit mechanisms in entorhinal cortex. *Neural Network*, 22, 1129–1138.
- Hasselmo, M. E. and Stern, C. E. (2006). Mechanisms underlying working memory for novel information. *Trends in Cognitive Sciences*, 10, 487–493.
- O'Keefe, J. and Recce, M. L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus*, 3, 317–330.

- Perry, E. K. and Perry, R. H. (1995). Acetylcholine and hallucinations: Disease-related compared to drug-induced alterations in human consciousness. *Brain and Cognition*, 28, 240–258.
- Sutton, R. S. and Barto, A. G. (1998). Reinforcement learning (adaptive computation and machine learning). Cambridge, MA, MIT Press.
- Tulving, E. (2002). Episodic memory: from mind to brain. *Annual Reviews Psychology*, 53, 1–25.
- Yoshida, M., Fransén, E. and Hasselmo, M. E. (2008). mGluR-dependent persistent firing in entorhinal cortex layer III neurons. *European Journal of Neuroscience*, 28, 1116–1126.
- Yoshida, M. and Hasselmo, M. E. (2009). Persistent firing supported by an intrinsic cellular mechanism in a component of the head direction system. *Journal of Neuroscience*, 29, 4945–52.
- Zilli, E. A. and Hasselmo, M. E. (2008). Analyses of Markov decision process structure regarding the possible strategic use of interacting memory systems. *Frontiers in Computational Neuroscience*, 2, 6.

SECTION II

Psychological processes



Birth of a Thought 3. Susan Aldworth (2007) etching and aquatint, 50 × 65 cm

Consciousness and the relation between implicit and explicit memory

Signy Sheldon and Morris Moscovitch

A commonly held assumption is that consciousness is a defining feature that distinguishes explicit memory (with conscious awareness) from implicit memory (without conscious awareness). Although early studies support this notion, recent evidence suggests that conscious and non-conscious memory systems may share crucial underlying processes. Here, we propose that one locus of interaction between some types of explicit and implicit memory may be the non-conscious processes associated with recollection, or detailed remembering, that are mediated by the hippocampus.

Consciousness has long been used to fractionate long-term memory into explicit and implicit memory; defined respectively as memory of a previous experience with and without conscious awareness (Schacter, 1987). Explicit memory is involved when one deliberately recovers past information and is often tested using recall and recognition tests. In recall tests, participants are asked to recall previously encountered items either freely or with the help of cues. In recognition tests, previously studied items and new items are presented together and the participants are asked to indicate which items were studied. In contrast, the hallmark of implicit memory tests is that the prior occurrence of an item or episode can bias the processing of a currently presented item without conscious awareness of retrieving the item from memory or, indeed, without any reference to the prior study episode at test.

An important method used in the study of implicit memory is *priming*. Priming is defined as processing facilitation caused by prior exposure to the information being tested and is typically measured by changes in performance on ostensibly non-memory tests. For example, exposing participants to a word at study and later asking them to perform a non-memory task such as deciding if a letter string is a word or not a word (lexical decision task) will result in faster and more accurate responses when the participants are exposed to the same word at study than if they are not.

Although distinguishing between explicit and implicit memory based on conscious awareness has a long history in psychology, current research has begun to question the validity of considering them as separable and non-interacting systems. In this chapter, we review the evidence that supports this distinction and the evidence that challenges it, concluding with the proposal of a new model for the interaction between the two memory systems.

The most compelling evidence for the distinction between explicit and implicit memory comes from early studies of patients with amnesia (for review see Moscovitch et al., 1993). These patients are characterized by an inability to remember past experiences consciously. This is evident when these patients perform tests of explicit memory such as free recall and recognition. For example, if these patients are given a list of words to study, they will show little memory for these words even on simple recognition tasks. Remarkably, in spite of this impairment, patients with amnesia have intact performance on tests of memory that do not require the conscious recovery of information from the learning episode, such as those assessing priming or procedural memory – the latter is a form of implicit memory that involves learning and executing perceptuo-motor skills such as riding a bike, mirror drawing or reading spatially transformed (inverted or backward) script. For example, in one study that used the stem-completion task, participants were asked to study a list of words, which included words such as “LANTERN”. Later, they were asked to think of words that could complete strings of three letters (e.g. LAN---) with no reference being made to the study episode. Implicit memory (priming) was assessed by the likelihood that a studied word was used complete the stems (e.g. LANTERN). In this study and many others like it, patients with amnesia demonstrated levels of priming comparable to those of healthy controls, and by implication, showed intact implicit memory (Graf et al., 1984) even though their explicit recall and recognition of the studied words were severely compromised. This led to the suggestion that the structures that are damaged in patients with amnesia, generally the medial temporal lobes (MTL) and more specifically the hippocampus, are crucially and exclusively involved in the conscious retrieval of explicit episodic memories (Tulving and Schacter, 1990).

Analogous dissociations of explicit and implicit memory have also been found using functional neuro-imaging techniques. These studies showed that retrieval of explicit memories was associated with *increased* activation in the medial temporal lobe (MTL) whereas implicit memory retrieval as assessed by priming was associated with *reduced* brain activity in regions implicated in perceptual (posterior cortex) and semantic (pre-frontal) processing of the stimulus (e.g. Schott et al., 2005; 2006).

Despite the support for using consciousness to distinguish between explicit and implicit memory, investigations have begun to challenge this assumption by

providing evidence that implicit and explicit memory, as traditionally defined, are not as separable as had once been assumed. Of note are those studies that have found that patients with MTL amnesia are also impaired on some tests of implicit memory. Many studies have reported that patients with amnesia have particular problems on implicit memory tests that involve relational or associative information in which two or more arbitrary features, attributes or stimuli are combined with each other. For example, using voice-specific priming, researchers found that, unlike healthy controls, patients with amnesia did not benefit when asked to identify spoken items that were presented in the same voice at study and test compared to when the voice differed from study to test (Schacter et al., 1995). In another study of implicit relational memory by Ryan and colleagues (2000), participants were asked to study scenes of real-world images. Then, at test, the participants viewed three types of scenes: new scenes, repeated old scenes, and manipulated old scenes in which positions of some objects had been altered. During test viewing, participant's eye movements were recorded as a measure of implicit memory since conscious awareness does not guide initial eye fixations. The researchers were specifically interested in the proportion of eye fixations to the region of the scene that was altered in the manipulated condition. Ryan and colleagues showed that the proportion of eye fixations to the manipulation was not related to explicit memory, indicating that eye movements directed to this area should be evident only if implicit memory is intact. Importantly, they found that neurologically intact participants had a greater number of fixations to the critical region of change in the manipulated conditions than the other two conditions. In contrast, however, patients with amnesia did not show this effect, indicating that their implicit memory processes related to this task were impaired. Importantly, the healthy controls were no more aware of the information that was manipulated than were the patients with amnesia.

Consistent with this evidence of impaired implicit memory in amnesia, are the findings that brain regions mediating explicit memory, specifically the hippocampus, are active during implicit memory tasks in neurologically intact people. For example, while in a PET scanner, participants were presented with words and asked to decide if the words belonged to a certain category. The words were then presented again, resulting in priming or a facilitation of the category decision. This study found that hippocampal activation at study accompanied word priming (Beauregard et al., 1998). A viable interpretation of these results, and other studies of its kind, is that the hippocampus, a brain region thought to be specialized for conscious episodic memory, may also be support performance on tests of implicit memory.

Following suit, more evidence began to emerge that showed specifically that explicit memory processes mediated by the hippocampus contributes to tasks

that do not require conscious recovery of past information, namely to implicit memory tasks. For example, Westmacott and colleagues (2003; 2004) showed that explicit memory, and recollection in particular, contributes to tests of semantic memory that do not require explicit reference to past episodes.

Before describing their tests, it is important to distinguish between two processes associated with explicit memory: recollection and familiarity. Recollection refers to the re-experiencing or reliving a past event in the mind such that the context in which a stimulus or event occurred can be recovered and re-created, a process that is thought to involve the hippocampus. In contrast, familiarity refers to recognition that a stimulus has been encountered in the study phase, but for which there is little or no information about the context of that encounter. Unlike recollection, familiarity is thought not to involve the hippocampus, but other regions within the medial temporal lobe, the likely candidate being the peri-rhinal cortex (Eichenbaum et al., 2007).

In one of their studies, Westmacott and colleagues (2003; 2004) gave healthy participants and patients with brain injuries or degeneration tests of fame judgment and speeded reading of famous names that were associated pre-experimentally with either high or low recollective experience (R). For example, Princess Diana is typically associated with a recollected memory of where the participant was and how he or she felt when learning about her fatal car accident. By contrast, George Bush senior, though equally familiar, did not evoke any particular memory or recollection in most participants. Not surprisingly, in healthy controls performance on tests of recall and recognition for these famous names (episodic based tests) was better for high R names than low R names. More interesting, however, was that their performance was also better (faster and more accurate) for the high R names than low R names on tests of fame judgment and speeded reading. These tests can be considered as conceptual and perceptual implicit memory tests, respectively, as they do not require conscious recovery of recollective experience. Most importantly, patients with medial temporal damage caused by focal lesions or degeneration did not show this benefit for high-R names on any of the tests, though other patients whose medial temporal lobes were relatively spared did show this benefit.

More recently, Hannula and Ranganth (2009) reported a study that also suggests that hippocampal activity can be associated with non-conscious, rapid, memory retrieval. While in an fMRI scanner, participants studied scenes on which faces were superimposed. The participants were then shown the scenes again with three faces and they were asked to indicate which face was previously associated with the presented scene. By tracking the participant's eye movements during this task, Hannula and Ranganath (2009) found that participants spent more time looking at correct faces than incorrect faces, even for trials in which the incorrect face

was explicitly selected and for trials in which the participants had no memory for which face matched the scene. In short, as in the previous study by Ryan and colleagues (2000), eye movements reflected implicit, but not explicit, memory. More interesting, however, was that the disproportionate viewing of the faces that match the scenes was correlated with hippocampal and surrounding medial temporal lobe regions activity, even in trials in which the participants could not explicitly recollect which face matched the scene being viewed at the time.

These studies, and many others, showed that processes thought to be exclusive to explicit memory also affect non-conscious processes associated with implicit memory. Based on these findings, Schacter and his colleagues (Schacter et al., 2004; 2007) proposed that there are two types of priming mechanisms, one implicated on associative tests of implicit memory, and the other, on non-associative tests. The latter type depends primarily on perceptual representational systems in the neo-cortex, whereas the former also involves the medial temporal lobes, which traditionally were considered explicit memory structures needed to form new associations, a function they proposed also applies to implicit memory tasks.

A related proposal advanced by Moscovitch (2008) has some elements in common with that of Schacter and his colleagues, but adds another dimension not considered by them. Rather than focus, as Schacter and his colleagues did, on associative or relational binding processes mediated by the MTL, Moscovitch emphasized recollective processes, which are dependent on the hippocampus and related structures, as the locus of interaction between explicit and implicit memory. Specifically, Moscovitch hypothesized that recollection involves two stages, a rapid, relatively automatic non-conscious stage, and a slower, more effortful conscious stage, which is associated primarily with explicit memory. It is at the first stage of recollection that explicit and implicit memory interact.

The two-stage model of recollection is based on key aspects of the Component Process Theory (Moscovitch, 1992). According to that model, both memory encoding and retrieval from the MTL are relatively automatic and obligatory. At encoding, the MTL, and hippocampus in particular, automatically encodes information that constitutes the conscious experience of an event. Encoding by the MTL, therefore, entails the formation of a memory trace which consists of a bound ensemble of MTL/hippocampal and neocortical neurons, the latter of which represent the conscious experience of an event. Thus, the memory trace represents both the to-be-remembered information about particular items and the conscious experience of an event that contains the context in which the items occurred. In short, the entire memory trace has all the elements needed for subsequent recollection. At retrieval, a cue that is either externally presented or internally generated reactivates the memory trace both automatically and obligatorily. Information recovered in this initial automatic stage is broadcast to all systems.

In so doing, on the one hand, it facilitates the processing of previously encountered stimuli which contributes to priming, and, on the other hand, it provides the materials needed for a slower process which enables the conscious recollection or re-experience of the initial event.

To recapitulate, the first stage of the recollection process is rapid and unconscious and the second stage is slower and conscious. The initial output from the first stage is broadcast widely and can contribute to performance on a variety of tasks, including priming and perhaps other tests of implicit memory. The second stage, which is slower and more effortful, is the one that is associated with conscious awareness of the retrieved information, thus making it explicit and available for guiding voluntary behavior.

To support this new theoretical proposal of implicit and explicit memory interactions, we recently conducted a series of studies to examine the relation between recollection and performance on two implicit memory tasks, lexical decision and word stem completion. Participants were asked to study a list of words, perform the required implicit memory task and then complete a recognition task ("Is this word old or new?") that also assessed if an old word was recollected or familiar via a procedure called the remember-know paradigm (Tulving, 1985; Figure 1, left panel). As predicted, we found that for both implicit tasks, words that were later shown to be recollected were associated with greater priming effects than were words which were not recollected but were merely recognized as familiar (Figure 1, right panel shows the results from the lexical decision experiment); the latter, familiar words showed the same amount of priming as words that had been studied but were not recognized (misses). These results suggest that only those explicit memory processes underlying recollection can benefit priming over and above that observed for words that were studied but not recognized, whereas those associated with familiarity, do not (Sheldon and Moscovitch, under review). These findings also suggest a reason why in some earlier studies performance on implicit tests were independent from performance on explicit ones and why in other studies they were not. The difference depended on whether the explicit test supported substantial levels of recollection, in which case priming would benefit from explicit memory, or whether explicit memory was based primarily on familiarity, in which case priming would not benefit from explicit memory.

In conclusion, evidence that implicit and explicit memory share memory representations and processes suggest these memory systems overlap, though the exact basis of this overlap is still being explored. Our proposal of a two stage model of recollection suggests that early processes underlying recollection mediated by the hippocampus may be one locus of interaction between conscious and non-conscious memory processes. By comparison, Schacter and colleagues' (2004; 2007) proposal places the emphasis on relational binding processing by the

hippocampus. Whereas all recollective processes are relational, not all relational processes are recollective, making the differences between the two theories eminently testable. Whichever view is favoured, these hypotheses demonstrate that, from a broader perspective, research on implicit and explicit memory is evolving from an enterprise that emphasized the separation and independence of the two systems, to one that seeks to understand their interaction, and the implication that interaction has for the relation between consciousness and memory.

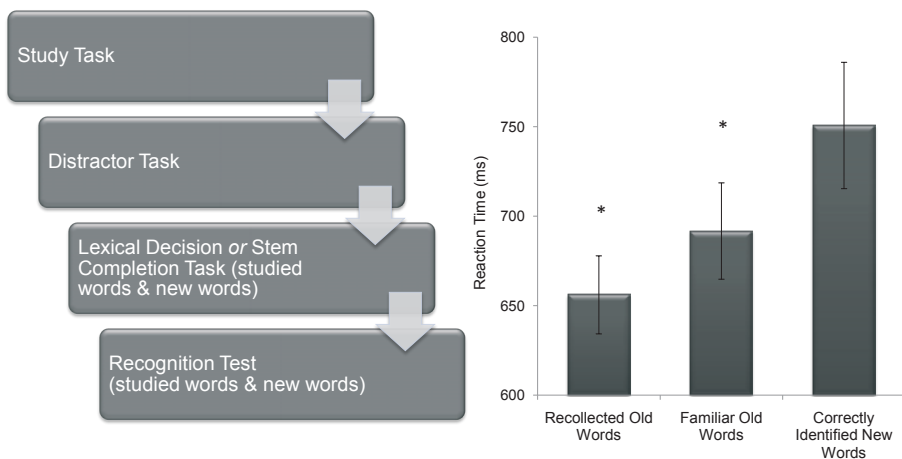


Figure 1. The left side depicts the procedure used in Sheldon and Moscovitch (under review) to assess the relation between recollection and priming. The right side of this figure depicts the results of the first half of this study: the mean lexical decision response time to words given a correct recognition response.

References

- Beauregard, M., Gold, D., Evans, A. C. and Chertkow, H. (1998). A role for the hippocampal formation in implicit memory: A 3-D PET study. *Neuroreport*, 9, 1867–1873.
- Eichenbaum, H., Yonelinas, A. R. and Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, 20, 123–152.
- Graf, P., Squire, L. R. and Mandler, G. (1984). The information that amnesic patients do not forget. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 164–178.
- Hannula, D. E. and Ranganath, C. (2009). The eyes have it: hippocampal activity predicts expression of memory in eye movements. *Neuron*, 63, 1–8.
- Moscovitch, M. (1992). Memory and working with memory: A component process model based on modules and central systems. *Journal of Cognitive Neuroscience*, 4, 257–267.
- Moscovitch, M., Vriezen, E. and Goshen-Gottstein, Y. (1993). Implicit tests of memory in patients with focal lesions or degenerative brain disorders. In F. Boller and J. Grafman (Eds.), *The handbook of neuropsychology*, Vol. 8. (pp. 133–173). Amsterdam: Elsevier Science Publishers.

- Moscovitch, M. (2008). The hippocampus as a “stupid” domain-specific module: Implications for theories of recent and remote memory, and of imagination. *Canadian Journal of Experimental Psychology*, 62, 62–79.
- Ryan, J. D., Althoff, R. R., Whitlow, S. and Cohen, N. J. (2000). Amnesia is a deficit in relational memory. *Psychological Science*, 11, 454–461.
- Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 501–518.
- Schacter, D. L., Church, B. A. and Bolton, E. (1995). Implicit memory in amnesic patients: Impairment of voice-specific priming. *Psychological Science*, 6, 20–25.
- Schacter, D. L., Dobbins, I. G. and Schnyer, D. M. (2004). Specificity of priming: A cognitive neuroscience perspective. *Nature Reviews Neuroscience*, 5, 853–862.
- Schacter, D. L., Wig, G. S. and Stevens, W. D. (2007). Reductions in cortical activity during priming. *Current Opinion in Neurobiology*, 17, 171–176.
- Schott, B. H., Henson, R. N., Richardson-Klavehn, A., Becker, C., Heinze, H. J. and Duzel, E. (2005). Redefining implicit and explicit memory: The functional neuroanatomy of priming, remembering, and control of retrieval. *Proceedings of the National Academy of Science of the USA*, 102, 1257–1262.
- Schott, B. H., Richardson-Klavehn, A., Henson, R. N., Becker, C., Heinze, H. J. and Duzel, E. (2006). Neuroanatomical dissociation of encoding processes related to priming and explicit memory. *Journal of Neuroscience*, 26, 792–800.
- Sheldon, S. and Moscovitch, M. (2010). Recollective Performance Advantages for Implicit Memory Tasks. (Under Review).
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, 26, 1–12.
- Tulving, E., and Schacter, D. L. (1990). Priming and human memory systems. *Science*, 247, 301–306.
- Westmacott, R. and Moscovitch, M. (2003). The contribution of autobiographical significance to semantic memory. *Memory and Cognition*, 31, 761–774.
- Westmacott, R., Black, S. E., Freedman, M. and Moscovitch, M. (2004). The contribution of autobiographical significance to semantic memory: evidence from Alzheimer’s disease, semantic dementia, and amnesia. *Neuropsychology*, 42, 25–48.

Two varieties of unconscious processes

Stan Franklin and Bernard J. Baars

Unconscious processes come in two varieties, the preconscious, whose contents may become conscious, and the never-conscious, whose contents may not. In this chapter we make use of Global Workspace Theory and its LIDA model to catalogue never-conscious and preconscious processes, and offer an explanation of the functional distinction. The LIDA/GWT model suggests that the functional distinction between never-conscious and preconscious processes derives from one of the major purposes of an agent's consciousness mechanism: that is to select the most salient portion of the agent's current situation to which to attend, in order for it to be broadcast globally, in order to choose the best next action.

Introduction

Cognitive processes¹ come in two varieties, conscious and unconscious, thought to be distinguished by the contents of conscious processes being broadly accessible throughout the nervous system (Baars, 2002). This conscious access hypothesis is consistent with a theory of conscious content being accessed throughout the thalamocortical core (Edelman and Tononi, 2000). A dynamical systems approach leads to the same hypothesis (Freeman, 2003). Philosophers have also discussed this hypothesis (Block, 2007). Finally, the conscious access hypothesis is a central tenet of Global Workspace Theory (see below, and Baars, 1988).

But unconscious processes also have their varieties. These are fundamental to our understanding of conscious events, because they provide the necessary comparison conditions. Unconscious brain processes often resemble conscious ones very closely, and it is the *differences* between them that provide clues about the nature of consciousness.

The contents of *preconscious processes* may remain unconscious or may become part of the conscious contents, while we will introduce the term "*Never-conscious*" to refer to those brain events that never come to consciousness – such

as automatisms, neuronal processes in the cerebellum, and long-term procedural memories. Dehaene et al. (2006) have suggested using *subliminal processing* for this purpose, but that term has been in common use for more than 50 years in the field of subliminal (sub-threshold) sensory perception. We therefore prefer to use “*Never-conscious*,” abbreviated as N-conscious.

The LIDA model of Global Workspace Theory (Franklin et al., 2005; Franklin and Patterson, 2006) (see below) enables us to draw the distinction between N-conscious vs. preconscious processes. The sections below will describe the LIDA model, discuss N-conscious and preconscious processes, and offer a case for the functionality of the distinction.

Global Workspace Theory

Global Workspace Theory (GWT, Baars, 1988) was originally conceived as a neuropsychological model of conscious and unconscious processes, but has been broadened into a higher-level theory of human cognitive processing, supported by empirical evidence (Baars, 2002), particularly when it is combined with the LIDA computational model of cognition. GWT views the nervous system as a distributed parallel system with many different specialized processes. Some coalitions of these processes enable an agent² to make sense of the sensory data coming from the current environmental situation. Other coalitions, incorporating the results of the processing of sensory data, compete for attention. The winner occupies what Baars calls a global workspace, the settled contents of which are broadcast to all other processes. In combination with other criteria, the broadcast contents of the global workspace are proposed to be conscious (see Baars, 1988). This conscious broadcast serves to recruit other, unconscious, processes to be used to select an action with which to deal with the current situation. GWT is therefore a theory of how consciousness functions within cognition. Unconscious contexts influence this competition for consciousness.

GWT is constantly being tested against new bodies of evidence. For example, any adequate theory must account for both conscious and unconscious brain conditions. The unconscious comparison conditions include:

1. Large brain regions that are constantly active, but never result in direct conscious output, such as the cerebellum and basal ganglia. The cerebellum, for example, may have as many neurons as the cerebral cortex. But cortex has traditionally been regarded as the source of conscious contents, while these other regions are not. Cortex has long been believed to provide the detailed contents of consciousness. A large number of studies show that lesions to

visual cortex results in selective types of blindness such as “blindsight,” face blindness (prosopagnosia), colour blindness or selective object blindness. Local or even global lesions in the cerebellum have no such effects, though the cerebellum may indirectly affect cortically-supported conscious contents. Since cortex and cerebellum have comparable numbers of neurons, it is clear that the sheer number of neurons involved in brain structures does not necessarily give rise to conscious experiences.

2. Other large regions of the cortex are *not* believed to give rise to conscious contents, notably the “dorsal visual stream,” which includes most of the parietal lobes (Goodale and Milner, 2004). It follows that merely being represented in cortex, even in regions sensitive to visual features, does not necessarily give rise to conscious contents. The dorsal stream may be best viewed as a “contextual” information stream that interacts with conscious objects, events, and scenes, all represented directly in the ventral stream. Dorsal visuotopic areas represent peripersonal space, object space, and the hand-eye coordination controlled by sensory and motor cortex, as in the case of reaching for an object or posting a letter in the slot of a mailbox. When the right parietal cortex is damaged, the left side of conscious visual space disappears from consciousness, though some unconscious processes continue. Such “parietal neglect” can be reversed through the irrigation of the left ear canal with cold water. The entire pattern of results suggests that the dorsal stream does not supply conscious object or event information by itself, but that it may provide a necessary condition for visual features and object represented in the ventral stream, including the MTL, to become conscious. This may sound like a subtle distinction, but it is absolutely indispensable. Basal brain nuclei that project to cortex for neuromodulation, for example, are crucial for the state of consciousness, without which we do not have specific experiences at all. Yet basal neuromodulating nuclei are not believed to directly support specific conscious contents.
3. The entire cortex is relatively less conscious for about six hours of the 24-hour cycle, during slow-wave sleep (SWS). Some researchers maintain that true unconsciousness occurs only during the so-called “DOWN” states of the delta wave during slow-wave sleep. It is certainly well established that some mental contents are often reported when people are awakened from SWS. Destexhe et al. (2007) have therefore suggested that even during SWS people may be fleetingly conscious during the UP states or peak of each slow oscillation. That question is under current study. If it is true, then totally unconscious events may only occur intermittently even during SWS, for a total of perhaps three hours per 24 hour cycle. Other evidence indicates that slow-wave sleep becomes deeper through the night, and that wider brain regions

decrease their metabolic activity over the hours of sleep. The greatest degree and duration of genuine unconsciousness may therefore occur in the latter half of the night.

4. It has long been known that novel skills are more available to consciousness than the identical behaviors after practice. Riding a bicycle is a common example. A great deal of evidence indicates that most cortical activity declines rapidly with the loss of conscious access to repeated skills (e.g. Schneider, 2009). Automatic (habitual) processes are believed to be handled by the basal ganglia and cerebellum when they become highly predictable. Cortical synaptic changes are no doubt involved as well. Once long-term synaptic changes have grown and stabilized it is possible that they are also less likely to give rise to conscious contents directly. However, this question is currently being studied by means of direct brain stimulation, and we can expect clearer results in the near future.
5. A number of experimental ways of manipulating conscious perceptual access have been carefully studied over half a century, including subliminal stimulation, visual backward masking, the attentional blink and selective attention. In many cases the unconscious comparison condition has been demonstrated to activate the same brain regions that are involved in the conscious condition, though to a lesser degree (e.g. Dehaene et al., 2006). Conscious contents, when compared to unconscious ones, generally evoke far more forward cortical activity in frontoparietal regions. Dehaene and colleagues take this to mean that global broadcasting takes place in the conscious case, arguably from the prefrontal regions using long axons. However, it is still possible that the sensory regions themselves may constitute global workspaces, using long fibres to send their information forward to frontoparietal cortex and across the corpus callosum. Baars and Franklin (2003) have suggested that the conscious broadcast (which may take 100 ms or multiples thereof) recruits areas involved in classical working memory, such as the dorso-lateral prefrontal cortex and the language areas of the speaking hemisphere. Classical working memory operates over 10–30 seconds, therefore much longer than the hypothesized global broadcast associated with GWT. Since brain oscillations are widely believed to carry signals among cortical regions, GWT may be associated with alpha or theta rhythms. It is generally thought that faster gamma activity may be multiplexed on such near-10 Hz waveforms. Alpha and theta activity may therefore have a pacing or grouping function to support high-gamma, long-distance resonance among cortical regions that together may be the physiological substrate of conscious moments.
6. A number of disease states lead to characteristic impairments of conscious states and experiences. These include certain epileptic seizures and epileptic

“behavioral automatisms,” loss of consciousness due to fainting, sleep disorders like sleepwalking, vegetative and minimally conscious states, analgesia and anesthesia, psychotic hallucinations and delusions, the dementias and so on. An adequate theory of conscious contents should be able to generate testable hypotheses about these conditions.

The LIDA model and its architecture

The LIDA model is a comprehensive, conceptual and computational³ model covering a large portion of human cognition (Franklin et al., 2007; Franklin and Patterson, 2006). Besides GWT, the model implements and fleshes out a number of psychological and neuropsychological theories including situated cognition (Varela et al., 1991), perceptual symbol systems (Barsalou, 1999), working memory (Baddeley and Hitch, 1974), memory by affordances⁴ (Glenberg, 1997), long-term working memory (Ericsson and Kintsch, 1995), and the H-CogAff architecture (Sloman, 1999). The comprehensive LIDA model includes a broad array of cognitive modules and processes, a database of which, including known possible neural correlates can be found online at <ccrg.cs.memphis.edu/tutorial/correlates.html>.

The LIDA cognitive cycle

The LIDA model and its ensuing architecture are grounded in the LIDA cognitive cycle. Every “autonomous agent” (Franklin and Graesser, 1997), human, animal, or artificial, must frequently sample (sense) its environment and select an appropriate response (action). Sophisticated agents, such as humans, process (make sense of) the input from such sampling in order to facilitate their decision making. Neuroscientists call this three-part process the action-perception cycle (Freeman, 2002). The agent’s “waking life” can be viewed as consisting of a continual sequence of these cognitive cycles. Each cycle constitutes a unit of sensing, processing and acting. A cognitive cycle can be thought of as a cognitive “moment.” Higher-level cognitive processes are composed of many of these cognitive cycles, so that each one can be considered to be a cognitive “atom.”

Just as atoms have inner structure, the LIDA model hypothesizes a rich inner structure for its cognitive cycles (Baars and Franklin, 2003; Franklin et al., 2005). During each cognitive cycle the LIDA agent first makes sense of (see below) its current situation as best as it can by updating its representation of both external and internal features of its world. By a competitive process to be described below,

it then decides what portion of the represented situation is most in need of attention. This portion is broadcast, making it the current content of consciousness, and enabling the agent to choose an appropriate action and execute it.

Figure 1 shows the process in more detail. It starts in the upper left corner and proceeds roughly clockwise.

The cycle begins with sensory stimuli from external and internal sources in the agent's environment. Low-level feature detectors in sensory memory begin the process of making sense of the incoming stimuli. These low-level features are passed on to perceptual associative memory where higher-level features, such as objects, categories, relations, events, situations, etc. are recognized. These entities, which have been recognized preconsciously, make up the percept that passes to the workspace, where a model of the agent's current situation is assembled. This percept serves as a cue to two forms of episodic memory, transient and declarative. Responses to the cue consist of local associations, that is, remembered events from these two memory systems that were associated with the various elements of the cue. In addition to the current percept, the workspace contains recent percepts and the models assembled from them that haven't yet decayed away.

A new model of the agent's current situation is assembled from the percepts, its associations, and the undecayed elements of the preceding model. This assembling process will typically require structure-building codelets⁵. These are small, special purpose processors, each of which has some particular type of structure it is designed to build. To fulfil their task these codelets may draw upon perceptual and sensory memory, to enable the recognition of relations and situations. The newly assembled model constitutes the agent's understanding of its current situation within its world. It has made sense of the incoming stimuli. All the contents of this workspace, including its current situational model, are preconscious, and may be selected to become conscious as we shall now see.

For an agent operating within a complex, dynamically changing environment, this current model may well be much too much for the agent to consider all at once in deciding what to do next. It needs to selectively attend to a portion of the model. Portions of the model therefore compete for attention. They take the form of coalitions of structures obtained from the current situational model. Such coalitions are formed by attention codelets, whose function is to bring certain structures to consciousness. Once one of the coalitions wins the competition, in effect, the agent has decided what to attend to – i.e., what to make conscious via the global workspace.

The purpose of this processing stage is to help the agent decide what to do next. To this end, a representation of the contents of the winning coalition is broadcast globally from the global workspace (hence the name Global Workspace Theory). Though the contents of this conscious broadcast are available globally,

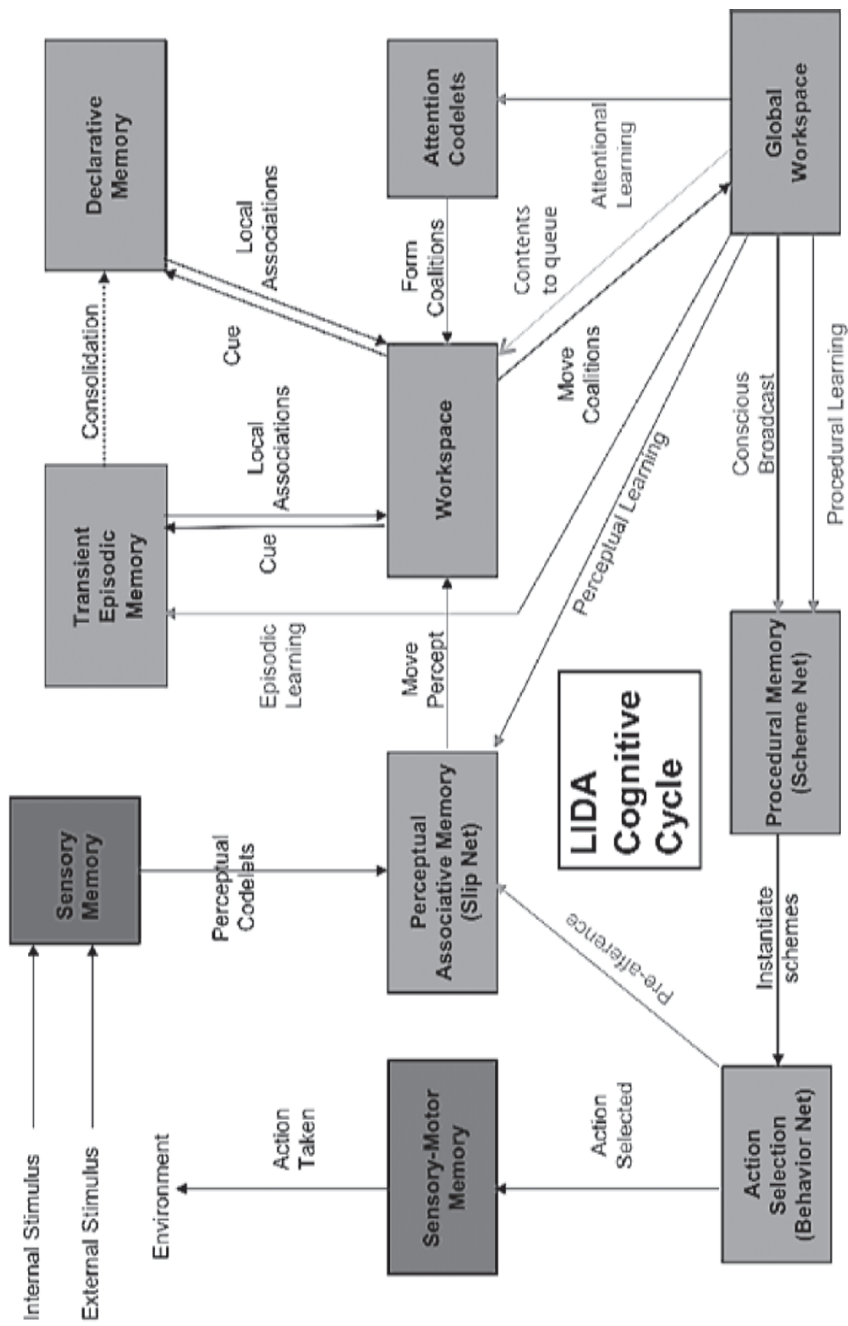


Figure 1. LIDA Cognitive Cycle Diagram. The boxes represent modules, not processes, and it is the processes that are preconscious or n-conscious. Some modules involve both kinds of processes.

a primary recipient is procedural memory, which stores templates of possible actions including their contexts and possible results. Procedural memory also stores an activation value for each template, which attempts to measure the likelihood of an action taken within its context producing the expected result. Templates whose contexts intersect sufficiently with the contents of the conscious broadcast instantiate copies of themselves with their variables specified to the current situation. Instantiated templates remaining from previous cycles may also continue to be available. These instantiations are passed to the action selection mechanism, which chooses a single action from one of these instantiations. The chosen action then goes to sensory-motor memory, from which it is executed by an appropriate algorithm. The action taken affects the environment, and the cycle is complete.

The LIDA model hypothesizes that all human cognitive processing is via a continuing iteration of such cognitive cycles. The unconscious elements of these cycles are proposed to occur asynchronously, with each cognitive cycle taking roughly 200–300 ms. These cycles cascade, that is, several cycles may have different processes running simultaneously in parallel. This cascading must, however respect the serial nature of conscious processes that are necessary to maintain the stable, coherent image of the world (Franklin, 2005; Merker, 2005). The cascading cycles, which partially overlap, allows a rate of cycling in humans of five to ten cycles per second. A cognitive “moment” is thus quite short! There is considerable evidence from cognitive psychology and neuroscience that is consistent with such cognitive cycling in humans (Massimini et al., 2005; Sigman and Dehaene, 2006; Uchida et al., 2006; Willis and Todorov, 2006).

Some of the strongest evidence comes from the study of alpha and theta rhythms, which are involved in visual immediate memory, and in hippocampal-neocortical integration of new episodic and semantic memories. Alpha and theta cycle in the 10-Hz neighbourhood and Steriade has indeed suggested that they are not categorically different. The same set of neurons may generate the slightly different pace of alpha and theta bands.

So far these questions are not yet settled, but the evidence appears consistent with our hypotheses.

Preconscious and N-conscious processes

Figure 1 illustrates the various cognitive modules and processes as hypothesized by the LIDA model, and the relationships between them. Beginning at the upper left, Sensory Memory consists of the initial representations of sensory input together with the products of the primary level feature detectors (Koch, 2004). Some of the earlier contents of Sensory Memory may be N-conscious, in that they can

never become conscious. The contents of Sensory Memory are sent to Sensory-Motor Memory to enable the execution of the action selected during each cognitive cycle. In the visual system, for example, the sensory-motor component of these contents travel along the dorsal stream (Goodale and Milner, 2004).

Several subsequent modules in the figure, Perceptual Associative Memory, the Workspace, and the two episodic memories, are concerned with making sense of the current situation. Any of the contents of their processes may be included in the winning coalition in the global workspace. Thus, all these contents are pre-conscious by the definitions given above.

The workings of the attention codelets and the global workspace (competition for consciousness and the conscious broadcast) never become conscious and are thus N-conscious. The same is true of the workings and the contents of Procedural Memory, Action Selection and Sensor-Motor Memory. All are N-conscious.

The LIDA/GWT model suggests that the functional distinction between N-conscious and preconscious process derives from one of the major roles of an agent's consciousness mechanism, to select the most salient portion of the agent's current situation to which to attend during each cognitive cycle. Only those processes whose contents might be part of that salient portion may be preconscious. Other critically important processes during any cognitive cycle serve many quite different functions, while their contents, if any, need never come to consciousness. For example, the action selection process during a single cognitive cycle is consciously mediated in that it requires momentary use of the current conscious contents. However the mechanism by which it selects the next action is entirely N-consciousness.

All higher-level mental activities, being multi-cyclic, must employ N-conscious, preconscious, and conscious processes during each of their cognitive cycles. During volition or conscious decision making, for example, a possible action will "pop into mind," that is, become conscious, perhaps in a fleeting visual image of the intended action outcome. This conscious thought is the result of both N-conscious and preconscious process operating during the cognitive cycle that produced the conscious event (Franklin, 2000).

Whether an unconscious cognitive process is N-conscious or preconscious depends entirely on its function. Does it carry content that might become part of a conscious broadcast? If so, it is preconscious by definition; if not, it is N-conscious.

Summary and conclusions

We have briefly sketched the GWT and LIDA models, and how they support information processing in autonomous agents, such as humans and animals. Consciousness is believed to involve momentary broadcasts of contents from one region or process in the brain, to be received by many others. Such broadcast events are believed to occur over approximately 100–200 ms bursts, consistent with our knowledge of alpha and theta activity. Distinctive contents of each burst may be multiplexed or modulated by faster gamma activity, which is often taken to mean 60–120 Hz. But “conscious bursts” are themselves embedded in cognitive cycles, which are unconscious. These cycles may be divided into preconscious elements, which can potentially provide conscious contents, and N-conscious, or “Never-conscious” elements, which are crucial to activating, and directing attention codelets to specific contents, but which never become conscious in themselves. There is a great deal of evidence that the brain supports both kinds of currently unconscious events.

The GWT-LIDA model therefore suggests there are two different kinds of unconscious events that occur quite fast and efficiently in any given cognitive cycles. The unconscious elements of such cycles may overlap in time, but the conscious ones face standard limited-capacity constraints and are therefore forced to occur serially.

Notes

1. In this paper we use the terms “process” and “content” in a strictly computational fashion. A process is thought of as an algorithm or a program as distinct from any content that might be produced by it.
2. This term refers to an autonomous agent in the sense of Franklin and Graesser (1997), which includes humans and most animals, as well as some software agents and mobile robots. Here we’re concerned mostly with human agents.
3. At this writing the LIDA model is only partially implemented. We claim it as a computational model since each of its modules and most of its processes have been designed for implementation.
4. Gibson (1979) introduced the term *affordance*, which is often interpreted to mean that information about the available uses of an object is evoked by the object itself. We are using it in the sense that the agent can derive such information from the object.
5. The term codelet refers generally to any small, special purpose processor or running piece of computer code.

References

- Baars, B. J. (1988). *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.
- Baars, J. (2002). The conscious access hypothesis: origins and recent evidence. *Trends in Cognitive Science*, 6, 47–52.
- Baars, B. J. and Franklin, S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Science*, 7, 166–172.
- Baddeley, A. D. and Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *The Psychology of Learning and Motivation* (pp. 47–89). New York: Academic Press.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577–609.
- Block, N. (2007). Consciousness, Accessibility and the Mesh between Psychology and Neuroscience. *Behavioral and Brain Sciences*, 30, 481–548.
- Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J. and Sergent, C. (2006). Conscious, pre-conscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences*, 10, 204–211.
- Destexhe, A., Hughes, S. W., Rudolph, M. and Crunelli, V. (2007). Are corticothalamic ‘up’ states fragments of wakefulness? *Trends in Neurosciences*, 30, 334–42.
- Edelman, G. M. and Tononi, G. (2000). *A Universe of Consciousness*. New York, NY: Basic Books.
- Ericsson, K. A. and Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102, 211–245.
- Franklin, S. (2000). Deliberation and voluntary action in ‘conscious’ software agents. *Neural Network World*, 10, 505–521.
- Franklin, S. (2005). Evolutionary pressures and a stable world for animals and robots: A commentary on Merker. *Consciousness and Cognition*, 14, 115–118.
- Franklin, S., Baars, B. J., Ramamurthy, U. and Ventura, M. (2005). The role of consciousness in memory. *Brains, Minds and Media*, 1, 1–38.
- Franklin, S. and Graesser, A. C. (1997). Is it an agent, or just a program? A taxonomy for autonomous agents. *Intelligent Agents III* (pp. 21–35). Berlin: Springer Verlag.
- Franklin, S. and Patterson, F. G. Jr. (2006). The LIDA Architecture: adding new modes of learning to an intelligent, autonomous, software agent. *IDPT-2006 Proceedings (Integrated Design and Process Technology)*: Society for Design and Process Science.
- Franklin, S., Ramamurthy, U., D’Mello, S. K., McCauley, L., Negatu, A. L. R. S., Silva, R. and Datla, V. (2007, November 9–11). LIDA: A computational model of global workspace theory and developmental learning. Paper presented at the AAAI Fall Symposium on AI and Consciousness: Theoretical Foundations and Current Approaches, Arlington, VA.
- Freeman, W. J. (2002). The limbic action-perception cycle controlling goal-directed animal behavior. *Neural Networks*, 3, 2249–2254.
- Freeman, W. J. (2003). Evidence from human scalp electroencephalograms of global chaotic itinerancy. *Chaos*, 13, 1067–1077.
- Gibson, J. J. and James, J. (1979). *The Ecological Approach to Visual Perception*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, 20, 1–19.
- Goodale, M. A. and Milner, D. (2004). *Sight Unseen*. Oxford: Oxford University Press.

- Koch, C. (2004). *The Quest for Consciousness: A neurobiological approach*. Englewood, Colorado: Roberts and Co.
- Massimini, M., Ferrarelli, F., Huber, R., Esser, Steve K., Singh, H. and Tononi, G. (2005). Break-down of cortical effective connectivity during sleep. *Science*, 309, 2228–2232.
- Merker, B. (2005). The liabilities of mobility: A selection pressure for the transition to consciousness in animal evolution. *Consciousness and Cognition*, 14, 89–114.
- Schneider, W. (2009). Automaticity and Consciousness. In W. Banks (Ed.), *Encyclopedia of Consciousness* (pp. 83–92). St. Louis: Elsevier.
- Sigman, M. and Dehaene, S. (2006). Dynamics of the central bottleneck: Dual-task and task uncertainty. *PLoS Biol* 4(7): e220. doi:10.1371/journal.pbio.0040220
- Solman, A. (1999). What Sort of Architecture is Required for a Human-like Agent? In M. Wooldridge and A. S. Rao (Eds.), *Foundations of Rational Agency* (pp. 35–52). Dordrecht, Netherlands: Kluwer Academic Publishers.
- Uchida, N., Kepecs, A. and Mainen, Z. F. (2006). Seeing at a glance, smelling in a whiff: Rapid forms of perceptual decision making. *Nature Reviews Neuroscience*, 7, 485–491.
- Varela, F. J., Thompson, E. and Rosch, E. (1991). *The Embodied Mind*. Cambridge, MA: MIT Press.
- Willis, J. and Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, 17, 592–599.

Operating characteristics and awareness

Colin W. G. Clifford, Justin A. Harris and Ehsan Arabzadeh

Although a large proportion of the human brain is dedicated to processing incoming sensory information, much of this processing is not accessible to consciousness. In an effort to understand what differentiates conscious from unconscious processing, we are concerned here with methodological issues in exploring the limits of unconscious processing. We first consider a conscious perceptual experience and establish an operating characteristic to disambiguate perceptual sensitivity from decision criterion. We then apply the same method to dissociate explicit measures of perception from implicit measures of sensory processing. Finally, we evaluate post-decision wagering as a proposed method for measuring awareness.

Disambiguating perceptual experience from decision criterion

How can I establish if you are conscious of something? For example, I might want to know if you were conscious of a very faint spot of light presented on a grey screen. The most natural approach would be to ask “Did you see the spot?” But if the spot were very faint you would probably be unsure whether you saw it or not. So how should I interpret whatever answer you give? If I present the spot on 10 trials then you might say “Yes, I think I saw something” on some and “No, I don’t think there was a spot” on others. You might say “Yes” on 6 trials and “No” on 4. Now if we run 10 more trials with a fainter spot maybe you would only say “Yes” twice. So far so good – your rate of reporting a conscious experience of the spot drops from 60% to 20% as the spot gets fainter. But now I tell you to say “Yes” only if you are absolutely sure that you saw the spot. I go back to presenting the spot at the original (higher) intensity but this time you only say “Yes” on 2 out of 10 trials. Evidently your response depends not only on your perceptual experience but also on a decision criterion as to whether that experience merits a “Yes” or “No” response. This is a problem if I just want to know whether you saw the spot. In this instance your criterion has changed as a result of an explicit instruction to be more conservative. However, the problem remains that I do not know how

conservative or liberal your criterion was in the first place. A liberal criterion might cause me to overestimate your perceptual sensitivity, while a conservative one would have the opposite effect.

To measure accurately your sensitivity to the spot of light, I need a “criterion-free” method so that I can factor out the effect of your decision criterion on your response. For example, I can include some “catch” trials where no spot is presented on the screen but I still ask if you saw the spot. By comparing your “hit rate” (the proportion of times that you correctly said “Yes” when the spot was presented) with your “false alarm rate” (the proportion of catch trials on which you erroneously reported seeing the spot), I can – subject to some assumptions – disambiguate the strength of your perceptual experience from your response criterion. We shall return to criterion-free methods and the assumptions underlying their analysis in the section on Operating Characteristics and Measures of Confidence in a Decision. We also direct the interested reader to the extensive literature on Signal Detection Theory (e.g. Wickens, 2002).

Dissociating explicit measures of perception from implicit measures of processing

Let us say that, armed with a criterion-free method for analysing your responses to presentations of the spot and intermingled catch trials, I find that your hit rate for detecting the spot at the fainter intensity is no greater than your false alarm rate. Can I safely conclude that you are not conscious of the spot at that intensity? Unfortunately, the answer is no. Statistically, one can infer that two distributions are significantly different, but it would be unsound to infer the absence of a difference on the basis of a lack of significance.

If we cannot establish rigorously that there is no conscious experience of a stimulus then how can we study *unconscious* processing? What we need is a means of dissociating our explicit measure of the strength of the perceptual experience elicited by a stimulus from an implicit measure of the processing of that stimulus. This implicit measure could be behavioural or physiological. An example of a behavioural measure would be the magnitude of an after-effect induced by an “invisible” stimulus (see Blake and He, 2005). In this case, the stimulus is undetected by explicit report (even with a criterion-free method), but there is a measurable impact of the stimulus on sensory processes as revealed indirectly by a subsequent change in explicit perceptual experience. A physiological measure might be the fMRI BOLD response of visual cortex to a stimulus flickering beyond the temporal resolution limit (e.g. Jiang et al., 2007). In this case, the

subject is incapable of distinguishing red-green flicker from an unchanging yellow field even though there is a clear effect on the blood-oxygen level in early visual cortex.

Consider the situation illustrated in Figure 1. Each panel plots hypothetical data comparing an explicit, criterion-free measure of perceptual experience, on the Y axis, with an implicit measure of the depth of processing (X axis). For example, one might compare (explicit) detection sensitivity for a pattern presented close to the spatial or temporal resolution limit of vision to the magnitude of an (implicit) after-effect induced following exposure to that pattern. In Figure 1A, the black circle located on the X axis indicates that there is significant processing of the stimulus, revealed by the implicit measure, even though the measure of explicit perception is not different from chance performance. There are many reports in the literature that claim such a pattern of results as being indicative of sensory processing without awareness, often in contrast to a condition where implicit and explicit measures are both well above zero. The underlying rationale would appear to be that, although a particular experimental manipulation (or natural accident) has only modestly reduced the depth of processing of a stimulus, the same manipulation has effectively abolished the observer's awareness of it.

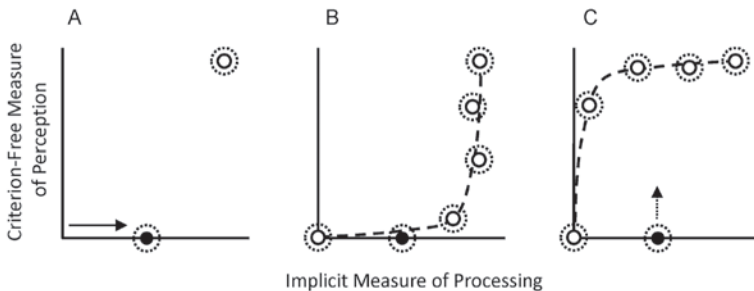


Figure 1. Operating characteristics to relate perception and processing.

Each panel plots hypothetical data comparing an explicit, criterion-free measure of perceptual experience (Y axis) with an implicit measure of the depth of processing (X axis). (A) The black circle located on the X axis indicates that there is significant processing of the stimulus, revealed by the implicit measure, even though the measure of explicit perception is not different from chance performance. The white circle represents a condition where implicit and explicit measures are both well above zero. (B) Parametric manipulation of stimulus visibility establishes an operating characteristic (dotted line) relating implicit and explicit measures in a baseline condition. (C) If the data point corresponding to the condition of interest lies below the operating characteristic established in the baseline condition, this is evidence for the involvement of unconscious processes.

Let us consider the situations illustrated in Figure 1B and 1C in the context of a concrete example. The black circle shows performance for conscious detection of a stimulus rendered invisible by, say, backward masking, plotted against the magnitude of an implicit measure (e.g. Clifford and Harris, 2005). Backward masking involves brief presentation of a visual stimulus, in this case an oriented grating, rapidly followed by a high contrast noise texture. The presentation of the noise texture masks perception of the grating, rendering it invisible. Despite being subjectively invisible, the masked grating affects the perceived orientation of an adjacent test grating that is itself unaffected by the mask. An implicit measure of the depth of processing of the invisible grating is thus provided by the magnitude of its effect on the perceived orientation of its unmasked neighbour.

The white circles represent hypothetical data in the absence of masking. Different white circles correspond to data from stimuli at different levels of contrast (note that the contrast level is not plotted on the graph, but we assume that the implicit and explicit measures are both monotonically increasing functions of contrast). Using contrast to parametrically manipulate stimulus visibility allows us to establish an “operating characteristic” relating implicit and explicit measures in the unmasked (baseline) condition. Specifically, decreasing grating contrast in the unmasked condition will decrease both its (explicit) detectability and the (implicit) effect it has on the perceived orientation of its neighbour in a characteristic manner. In each of Figures 1B and 1C this operating characteristic is a monotonically increasing function (the dotted line in both figures); this monotonic relationship between implicit and explicit measures follows from the assumption that each is related monotonically to stimulus contrast.

In Figure 1B, the magnitude of the implicit measure saturates at a lower contrast than the explicit measure; vice versa in Figure 1C. A common interpretation of data like those in Figure 1B is that the black data point indicates unconscious processing. However, the flaw in this reasoning is made clear by referring to Figure 1C. To say that any data point lying on the operating characteristic in Figure 1B was evidence of “processing without awareness” would seem to us no more sensible than to maintain that a data point lying on the operating characteristic in Figure 1C was evidence of “awareness without processing”! Instead, if we wish to claim evidence for a dissociation in the effect of masking on implicit processing versus explicit awareness, the data point from the backward masking condition should not lie on the operating characteristic in the unmasked condition.

If the data point corresponding to the masked condition lies below the operating characteristic established in the absence of masking, as in Figure 1C, this is evidence that masking has a disproportionately greater effect upon

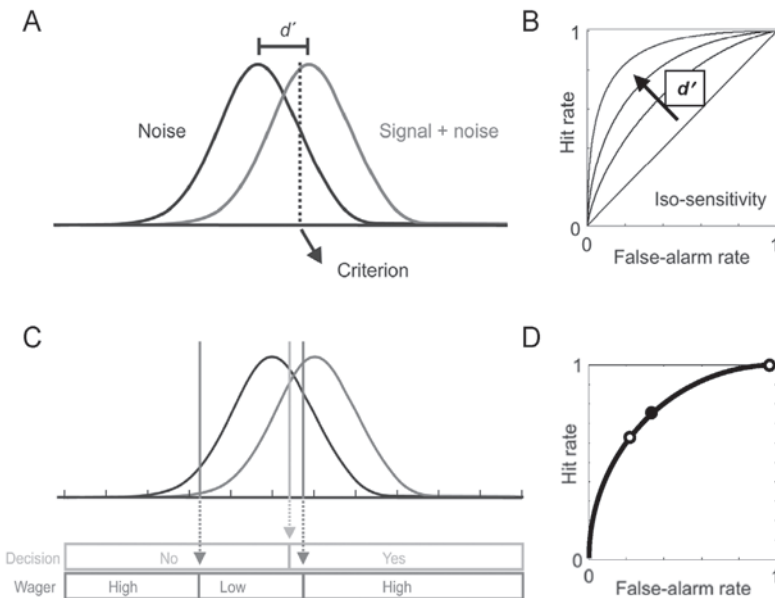
awareness than upon our measure of processing. This implies the involvement of unconscious processes in mediating the pattern-specific effect we are using as our implicit measure. Such a pattern of results could occur if, for example, masking predominantly affected higher-level, conscious processes subsequent to lower-level, unconscious processes that mediate the pattern-specific effect. In the example given above, this would correspond to a situation in which backward masking left lower-level orientation-specific interactions between neighbouring gratings largely unaffected but intervened to block access to higher-level, conscious processes. If instead the effect were mediated purely by conscious processes, then the pattern-specific effect would not show greater immunity to masking than the corresponding perceptual experience. In this case the data point corresponding to the masked condition would lie on the same operating characteristic as the unmasked data.

It is important to note that a dissociation between processing and awareness of the type illustrated in Figure 1C does not require that perceptual performance is “at chance” in any condition. Even if there were evidence of significant perceptual awareness in all conditions it might still be possible to show that the data from the condition of interest are not consistent with the established operating characteristic. However, with real experimental data, it will not always be easy to tell by eye whether or not a given data point is consistent with a particular operating characteristic. Fortunately, formal methods exist that can be used in such situations to test whether all the data can be captured by a single operating characteristic. If they can, then the data provide no evidence for processing without awareness. If, instead, different characteristics are warranted to account for the masked and unmasked data, then that indicates the operation of unconscious processes in the masked condition. An exposition of such methods is beyond the scope of this chapter. However, we direct the interested reader to the literature on State-Trace Analysis (Bamber, 1979; Newell and Dunn, 2008) and model comparison methods (Burnham and Anderson, 1998).

Operating characteristics and measures of confidence in a decision

It has recently been proposed that a direct measure of awareness in a perceptual task can be obtained by asking the subject to make a wager on the outcome of their decision (Persaud et al., 2007). For example, after making a “Yes-No” judgment as to whether a stimulus was present, the subject is then required to bet on the correctness of that perceptual judgment. The subject is given the opportunity of wagering high or low, depending on how confident they are in the correctness

of their decision. The claim that post-decision wagering objectively measures awareness has been criticized on a number of grounds. A post-decision wager is a meta-cognitive judgment on the correctness of a perceptual decision, so it is not clear in what sense this could be considered a “direct” measure of awareness (Seth, 2008). Indeed, in the literature on behavioural economics, wagering is known to be affected by loss aversion, whereby to the subject the prospect of losing a given sum in a wager can only be offset by the prospect of winning an even greater sum of money (Kahneman and Tversky, 1979; Schurger and Sher, 2008). Thus, we have argued previously that failure to optimize ones earnings from post-decision wagering is not a reliable indication of lack of awareness of the sensory evidence upon which the original decision was based (Clifford et al., 2008a, b). Indeed, we have demonstrated that an empirical data set used as an example of how to establish perception without awareness with post-decision wagering (Persaud et al., 2007) can in fact be described by a single operating characteristic. Specifically, we showed that the failure of the subject to optimize his earnings from post-decision wagering was consistent simply with a reluctance of the subject to wager high on a decision when the sensory evidence upon which that decision was based was below a certain criterion level (Figure 2). In other words, the subject’s behaviour was consistent with him using a more conservative criterion for betting high than for responding “Yes” to the presence of the stimulus, but all in full awareness of the sensory evidence (Clifford et al., 2008a).



One point we did not address explicitly in our critique of the study by Persaud et al. (2007) was their use and interpretation of a Type II signal detection measure (Galvin et al., 2003). In Signal Detection Theory, Type I measures quantify the subject's ability to perform a perceptual task while Type II measures quantify the ability to distinguish correct from incorrect decisions on the perceptual task. If one is unaware of the information upon which one is basing one's perceptual decisions, then there should be no relationship between the accuracy of a decision and measures of subjective confidence in that decision (Dienes et al., 1995). Persaud et al. (2007) found that the post-decision wagering performance of their subject yielded a Type II measure that was not significantly different from chance. However, as we have already mentioned, failure of a measure to demonstrate significant awareness is not a demonstration of lack of awareness. Moreover, it should be noted that the particular measure of Type II performance used by Persaud et al. (2007) can vary with changes in response bias that affect the relative frequency of high and low wagers (Evans and Azzopardi, 2007). Indeed, as noted above, we found that the value of the Type II measure computed by Persaud et al. (2007) was entirely consistent with adoption of a conservative, loss-averse criterion for wagering high on the outcome of a perceptual decision in full awareness of the sensory evidence upon which that decision was based.

Conclusion

Neuroscience is engaging ever more sophisticated techniques in the study of consciousness. In this chapter we have focussed on fundamental methodological issues in establishing evidence of unconscious processing. The application of Type I signal detection theory ensures that changes in perceptual experience

Figure 2. Signal Detection Theory.

(A) In a Yes-No detection task, Signal Detection Theory assumes that an observer will respond positively when the sensory evidence for the presence of the target stimulus exceeds a criterion level. (B) The subjective strength (d') of the stimulus determines the shape of the operating characteristic, whereas variations in criterion level determine the location on that operating characteristic. (C) Decision and wagering criteria for a hypothetical observer (Clifford et al., 2008a) designed to reproduce the performance levels reported by Persaud et al. (2007). Each wager is based upon the same sensory evidence as the preceding decision and differs only in criterion level. (D) The data used by Persaud et al. (2007) as an example of how to establish perception without awareness with post-decision wagering can thus in fact be described by a single operating characteristic.

are not confounded with changes in decision criterion. We have advocated the measurement of operating characteristics that relate explicit reports of perceptual experience to an implicit measure of neural processing. This implicit measure could itself be behavioural, such as the magnitude of an after effect, or physiological, such as the cortical fMRI BOLD response. Demonstrating that the operating characteristic between neural activity and conscious perception varies as a function of stimulus presentation (e.g. masked versus unmasked) is evidence of unconscious processing. Finally, we urge caution in the use of post-decision wagering and in particular the interpretation of Type II signal detection theory measures as evidence of unconscious processing.

Acknowledgments

C. C. is supported by an Australian Research Fellowship and E. A. by an Australian Post-doctoral Fellowship.

References

- Bamber, D. (1979). State-trace analysis: A method of testing simple theories of causation. *Journal of Mathematical Psychology*, 19, 137–181.
- Blake, R. and He, S. (2005). Adaptation as a tool for probing the neural correlates of visual awareness: progress and precautions. In C. W. G. Clifford and G. Rhodes (Eds.), *Fitting the Mind to the World: Aftereffects in High-Level Vision* (pp. 281–307). Oxford: Oxford University Press.
- Burnham, K. P. and Anderson, D. R. (1998). *Model selection and multi-model inference*. New York, NY: Springer.
- Clifford, C. W. G. and Harris, J. A. (2005). Contextual modulation outside of awareness. *Current Biology*, 15, 574–578.
- Clifford, C. W. G., Arabzadeh, E. and Harris, J. A. (2008a). Getting technical about awareness. *Trends in Cognitive Sciences*, 12, 54–58.
- Clifford, C. W. G., Arabzadeh, E. and Harris, J. A. (2008b). A good bet to measure awareness? *Trends in Cognitive Sciences*, 12, 210.
- Dienes, Z., Altmann, G., Kwan, L. and Goode, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1322–1338.
- Evans, S. and Azzopardi, P. (2007). Evaluation of a ‘bias-free’ measure of awareness. *Spatial Vision*, 20, 61–77.
- Galvin, S. J., Podd, J. V., Drga, V. and Whitmore, J. (2003). Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychonomic Bulletin and Review*, 10, 843–876.

- Jiang, Y., Zhou, K. and He, S. (2007). Human visual cortex responds to invisible chromatic flicker. *Nature Neuroscience*, 10, 657–662.
- Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291.
- Newell, B. R. and Dunn, J. C. (2008). Dimensions in data: Testing psychological models using state-trace analysis. *Trends in Cognitive Sciences*, 12, 285–290.
- Persaud, N., McLeod, P. and Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nature Neuroscience*, 10, 257–261.
- Schurger, A. and Sher, S. (2008). Awareness, loss aversion and post-decision wagering. *Trends in Cognitive Sciences*, 12, 209–210.
- Seth, A. K. (2008). Post-decision wagering measures metacognitive content, not sensory consciousness. *Consciousness and Cognition*, 17, 981–983.
- Wickens, T. D. (2002). *Elementary signal detection theory*. New York, NY: Oxford University Press.

Noise in the brain, decision-making, determinism, free will, and consciousness

Edmund T. Rolls

Noise caused by randomness in the spiking times of neurons in the brain has a number of advantages, including contributing to probabilistic decision-making. However, noise results in the brain operating effectively as a non-deterministic system, which has implications for free will. Noise also results in decisions being taken probabilistically between the reasoning system and the implicit reward system. I propose that free will can be used to describe the operation of the reasoning system, and that consciousness is a property of a reasoning system that must use higher order syntactic thoughts (HOSTs) to correct its first order thoughts. When the implicit system takes a decision, we may confabulate a reason for the decision, and in that case the feeling of free will may be an illusion.

1. Introduction

In this paper I consider the implications of noise in the brain for determinism, for freewill, for decisions between the conscious and implicit (unconscious) routes to action, and for monitoring processes in the brain, in the context of a theory of emotion (Rolls, 2005a), of consciousness (Rolls, 2004; 2007a; 2008a; in press), and of noise in the brain (Rolls and Deco, 2010).

2. A higher order syntactic theory of consciousness

In a reasoned plan with several steps to the plan, if there is an error, how do we know which step has the error? There is a credit assignment problem here, for the occurrence of an error does not tell us which step had a fault. I argue that in this situation, thoughts about the steps of the plan, that is thoughts about thoughts, namely “higher order thoughts”, can help us to detect which was the weak step in the plan, with perhaps weak premises, so that we can correct the plan and try again. I argue that having thoughts about our previous thoughts, reflecting

on them, may be a computational process that, when it occurs, is accompanied by feelings of consciousness, that it feels like something (Rolls, 1997a, b; 2003; 2004; 2005b; 2007a, b; 2008a; in press). This general approach to the phenomenal aspects of consciousness is shared by a number of philosophers and scientists (Rosenthal, 1990; 1993; 2004; 2005; Weiskrantz, 1997). I then go on to argue that when this 'higher order syntactic thought' (HOST) brain processor is actively processing sensory states or emotions (about which we may sometimes need to reason), then conscious feelings about these sensory states or emotions, called qualia, are present. I argue that this reasoning (i.e. rational) explicit (conscious) system is propositional and uses syntactic binding of symbols.

I contrast this rational or reasoning, explicit, system with an implicit (unconscious) system that uses gene-specified goals (e.g. food reward when hungry, water when thirsty, social interaction (Rolls, 2005a)) for actions, and can perform arbitrary actions to obtain the genotypically defined goals. The explicit system, because of the reasoning, can perform multiple step planning which might lead to goals that are in the interest of the phenotype but not of the genotype (e.g. not having children in order to devote oneself to the arts, philosophy, or science). In this sense, we can speak of the *selfish genotype* vs the *selfish phenotype*. If the implicit system takes a decision and performs an action, confabulation may occur, in that a verbal account may be given by the explicit (conscious) system of why the action was performed, and this may not be related at all to the environmental event that actually triggered the action (Gazzaniga and LeDoux, 1978; Gazzaniga, 1988; 1995; LeDoux, 2008; Rolls et al., 1994; Rolls, 2005a; 2008a; in press).

3. Noise in the brain and determinism

There is internal noise (i.e. random effects) in the brain caused by the timing of the action potentials of each neuron, which occur at approximately random times for a given mean firing rate, that is, approximate a Poisson process (Rolls and Deco, 2010). The underlying cause of this randomness is that the membrane potential is close to the threshold of firing, so that small changes in the input to a neuron can cause it to suddenly emit an action potential. The changes in input can arise from noise in the environment or in the sensors (such as the eye), but also are related to internal probabilistic opening of ion channels in the neuronal membrane and in synapses (Faisal et al., 2008; Rolls and Deco, 2010). This noise is in fact helpful, in that it breaks the symmetry of many decision-making and memory recall processes in the brain, allowing probabilistic decision-making which is adaptive, and creative thought because each thought is different (Rolls and Deco, 2010). To provide some insight into the probabilistic decision-making mechanism, if we

have equal inputs to an integrate-and-fire decision-making network, then on an individual trial, more spiking in the neurons in attractor 1 which represents decision 1 might occur because of randomness in the spiking of the neurons; and on another trial, more spiking in the neurons in attractor 2 which represents decision 2 might occur because of randomness in the spiking of the neurons (Rolls and Deco, 2010) (see Figure 1).

Given that the brain operates with some degree of randomness due to the statistical fluctuations produced by the random spiking times of neurons, brain function is to some extent non-deterministic, as defined in terms of these statistical fluctuations. That is, the behaviour of the system, and of the individual, can vary from trial to trial based on these statistical fluctuations, in ways that are described by Rolls and Deco (2010). Given that each neuron has this randomness, and that there are sufficiently small numbers of synapses on each of the neurons in each network (between a few thousand and 20,000) that these statistical fluctuations are not smoothed out (Rolls and Deco, 2010), and that there are a number of different networks involved in typical thoughts and actions each one of which may behave probabilistically, and with 10^{11} neurons in the brain each with this number of synapses, the system has so many degrees of freedom that it operates effectively as a non-deterministic system. (Philosophers may wish to argue about different senses of the term deterministic, but it is being used here in a precise, scientific, and quantitative way, which has been clearly defined (Rolls and Deco, 2010).)

4. Noise in the brain and free will

This raises the issue of free will in decision-making. One factor in this is that as I have suggested the brain operates effectively non-deterministically. But there is what I think is an even more interesting factor, which I believe gets at the heart of a useful way to think about free will. It comes from considering the operation of the implicit and explicit systems in the brain, as follows.

Do we have free will when both the implicit and the explicit systems have made the choice? Free will would in my view (Rolls, 2005a; 2008a, b; in press) involve the use of language to check many moves ahead on a number of possible series of actions and their outcomes, and then with this information to make a choice from the likely outcomes of different possible series of actions. (If in contrast choices were made only on the basis of the reinforcement value of immediately available stimuli, without the arbitrary syntactic symbol manipulation made possible by language, then the choice strategy would be much more limited, and we might not want to use the term free will, as all the consequences of those

actions would not have been computed.) It is suggested that when this type of reflective, conscious, information processing is occurring and leading to action, the system performing this processing and producing the action would have to believe that it could cause the action, for otherwise inconsistencies would arise (with for example the reasoning system feeling out of control if it was clear that the action was being initiated by the implicit system), and the rational system might no longer try to initiate action. This belief held by the rational system may partly underlie the feeling of free will. At other times, when other brain modules are initiating actions (in the implicit systems), the conscious processor (the explicit, reasoning, system) may confabulate and believe that it caused the action, or at least give an account (possibly wrong) of why the action was initiated. The fact that the conscious rational processor may have the belief, even when the implicit system has taken the decision, that it, the rational system, initiated the action may arise as a property of it being inconsistent for a rational system that can take overall control using conscious syntactic processing to believe that it was over-ridden by another system. This may be the underlying computational reason why confabulation occurs.

The implication of this argument is that a good use of the term free will is when the term refers to the operation of the rational, planning, explicit (conscious) system that can think many moves ahead, and choose from a number of such computations the multistep strategy that best optimises the goals of the explicit system with long-term goals. When on the other hand our implicit system has taken a decision, and we confabulate a spurious account with our explicit system, and pronounce that we took the decision for such and such a (confabulated) reason of our own “free will”, then I want to say that the feeling of free will was an illusion (Rolls, 2005a; in press).

5. Monitoring and consciousness

An attractor network in the brain with positive feedback implemented by excitatory recurrent collateral connections between the neurons can implement decision-making (Deco and Rolls, 2006; Rolls and Deco, 2010; Wang, 2002). As explained in detail elsewhere (Rolls and Deco, 2010), noise caused by the random spiking of the neurons can influence the decision-making. If the external evidence is consistent with the decision taken (which has been influenced by the noisy neuronal firing times), then the firing rates in the winning attractor are supported by the external evidence, and become especially high. If the external evidence is contrary to the noise-influenced decision, then the firing rates of the neurons in the winning attractor are not supported by the external evidence, and

are lower than expected. In this way the confidence in a decision is reflected in, and encoded by, the firing rates of the neurons in the winning attractor population of neurons (Rolls et al., 2009; 2010).

If we now add a second attractor network to read the firing rates from the first decision-making network, the second attractor network can take a decision based on the confidence expressed in the firing rates in the first network (Insabato et al., 2010). The second attractor network allows decisions to be made about whether to change the decision made by the first network, and for example abort the trial or strategy (see Figure 1). The second network, the confidence decision network, is in effect monitoring the decisions taken by the first network, and can cause a change of strategy or behaviour if the assessment of the decision taken by the first network does not seem a confident decision. This is described in detail elsewhere (Insabato et al., 2010), but Figure 1 shows the simple system of two attractor networks that enables confidence-based (second level) decisions to be made, by monitoring the output of the first, decision-making, network.

Now this is the type of description, and language used, to describe 'monitoring' functions, taken to be high level cognitive processes, possibly related to consciousness (Block, 1995; Lycan, 1997). For example, in an experiment performed by Hampton (2001) (experiment 3), a monkey had to remember a picture over a delay. He was then given a choice of a 'test flag', in which case he would be allowed to choose from one of four pictures the one seen before the delay, and if correct earn a large reward (a peanut). If he was not sure that he remembered the first picture, he could choose an 'escape flag', to start another trial. With longer delays, when memory strength might be lower partly due to noise in the system, and confidence therefore in the memory on some trials might be lower, the monkey was more likely to choose the escape flag. The experiment is described as showing that the monkey is thinking about his own memory, that is, is a case of meta-memory, which may be related to consciousness (Heyes, 2008). However, the decision about whether to escape from a trial can be taken just by adding a second decision network to the first decision network. Thus we can account for what seem like complex cognitive phenomena with a simple system of two attractor decision-making networks (Figure 1) (Rolls and Deco, 2010).

The implication is that some types of 'self-monitoring' can be accounted for by simple, two-attractor network, computational processes. But what of more complex 'self-monitoring', such as is described as occurring in a commentary that might be based on reflection on previous events, and appears to be closely related to consciousness (Weiskrantz, 1997). This approach has been developed into my higher order syntactic theory (HOST) of consciousness (Rolls 1997a, b; 2003; 2004; 2005b; 2007a, b; 2008a; in press), in which there is a credit assignment problem if a multi-step reasoned plan fails, and it may be unclear which step failed.

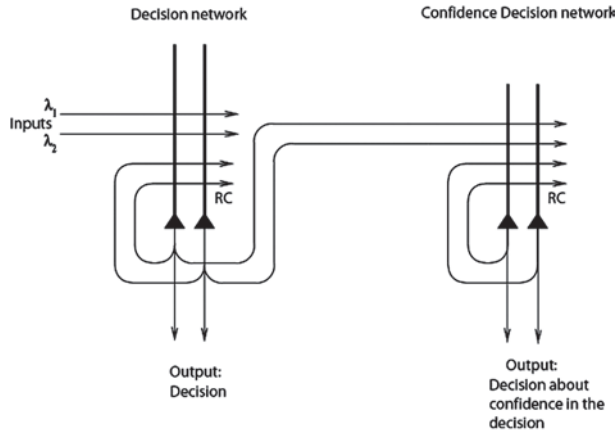


Figure 1. Network architecture for decisions about confidence estimates. The first network is a decision-making network, and its outputs are sent to a second network that makes decisions based on the firing rates from the first network, which reflect the decision confidence. In the first network, high firing of population D1 represents decision 1, and high firing of population D2 represents decision 2. The second network is a confidence decision network, and receives inputs from the first network. The confidence network has two selective pools of neurons, one of which (C) responds to represent confidence in the decision, and the other of which responds when there is little or a lack of confidence in the decision (LC). Each selective pool has strong recurrent connections, while the connections between the two selective pools are weak. All other connections are set to the default value 1. All neurons in the network receive an external input (λ_{ext}) emulating the spontaneous firing of neurons in other cortical areas. Pools D1 and D2 receive also a stimulus-related input (respectively λ_1 and λ_2) between which a decision is taken by the first (decision-making) network. RC – recurrent collateral connections. (After Insabato et al., 2010.)

Such plans are described as syntactic as there are symbols at each stage that must be linked together with the syntactic relationships between the symbols specified, but kept separate across stages of the plan. It is suggested that in this situation being able to have higher order syntactic thoughts will enable one to think and reason about the first order plan, and detect which steps are likely to be at fault.

Now this type of ‘self-monitoring’ is much more complex, as it requires syntax. The thrust of the argument is that some types of ‘self-monitoring’ are computationally simple, for example in decisions made based on confidence in a first decision (Insabato et al., 2010; Rolls and Deco, 2010), and may have little to do with consciousness; whereas higher order thought processes are very different in terms of the type of syntactic computation required, and may be more closely related to consciousness (Rolls 1997a, b; 2003; 2004; 2005b; 2007a, b; 2008a; in press).

In conclusion, it is argued that complex issues such as how decisions are made between the conscious and non-conscious systems in the brain, about free will, and even about what consciousness is and what its functions are, are illuminated by advances in understanding decision processes in the brain, and how noise affects these, at a computational neuroscience level (Rolls, 2008b; Rolls and Deco, 2010), and how different processing systems in the brain have had different evolutionary origins and perform different functions leading to a contrast between the selfish genotype vs the selfish phenotype (Rolls, 2005a; in press).

References

- Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences* 18, 227–247.
- Deco, G. and Rolls, E. T. (2006). Decision-making and Weber's law, a neurophysiological model. *European Journal of Neuroscience*, 24, 901–916.
- Faisal, A. A., Selen, L. P. and Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience*, 9, 292–303.
- Gazzaniga, M. S. (1988). Brain modularity, towards a philosophy of conscious experience. In A. J. Marcel and E. Bisiach (Eds.), *Consciousness in Contemporary Science* (pp. 218–238). Oxford: Oxford University Press.
- Gazzaniga, M. S. (1995). Consciousness and the cerebral hemispheres. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (pp. 1392–1400). Cambridge, MA: MIT Press.
- Gazzaniga, M. S. and LeDoux, J. (1978). *The Integrated Mind*. New York, NY: Plenum Press.
- Hampton, R. R. (2001). Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences of the USA.*, 98, 5359–5362.
- Heyes, C. M. (2008). Beast machines? Questions of animal consciousness. In L. Weiskrantz and M. Davies (Eds.), *Frontiers of Consciousness* (pp. 259–274). Oxford, Oxford University Press.
- Insabato, A., Pannunzi, M., Rolls, E. T. and Deco, G. (2010). Confidence-related decision-making. *Journal of Neurophysiology*, 104, doi:10.1152/jn.01068.2009.
- LeDoux, J. (2008). Emotional coloration of consciousness, how feelings come about. In L. Weiskrantz and M. Davies (Eds.), *Frontiers of Consciousness* (pp. 69–130). Oxford: Oxford University Press.
- Lycan, W. G. (1997). Consciousness as internal monitoring. In N. Block, O. Flanagan and G. Guzeldere (Eds.), *The Nature of Consciousness, Philosophical Debates* (pp. 755–771). Cambridge, MA: MIT Press.
- Rolls, E. T. (1997a). Consciousness in neural networks? *Neural Networks*, 10, 1227–1240.
- Rolls, E. T. (1997b). Brain mechanisms of vision, memory, and consciousness. In M. Ito, Y. Miyashita and E. T. Rolls (Eds.), *Cognition, Computation, and Consciousness* (pp. 81–120). Oxford: Oxford University Press.
- Rolls, E. T. (2003). Consciousness absent and present, a neurophysiological exploration. *Progress in Brain Research*, 144, 95–106.
- Rolls, E. T. (2004). A higher order syntactic thought (HOST). theory of consciousness. In R. J. Gennaro (Ed.), *Higher-Order Theories of Consciousness, An Anthology* (pp. 137–172). Amsterdam: John Benjamins.

- Rolls, E. T. (2005a). *Emotion Explained*. Oxford: Oxford University Press.
- Rolls, E. T. (2005b). Consciousness absent or present, a neurophysiological exploration of masking. In H. Ogmen and B. G. Breitmeyer (Eds.), *The first half second, the microgenesis and temporal dynamics of unconscious and conscious visual processes* (pp. 89–108). Cambridge, MA: MIT Press.
- Rolls, E. T. (2007a). A computational neuroscience approach to consciousness. *Neural Networks*, 20, 962–982.
- Rolls, E. T. (2007b). The affective neuroscience of consciousness, higher order linguistic thoughts, dual routes to emotion and action, and consciousness. In P. Zelazo, M. Moscovitch and E. Thompson (Eds.), *Cambridge Handbook of Consciousness* (pp. 831–859). Cambridge: Cambridge University Press.
- Rolls, E. T. (2008a). Emotion, higher order syntactic thoughts, and consciousness. In L. Weiskrantz and M. K. Davies (Eds.), *Frontiers of Consciousness* (pp. 131–167). Oxford: Oxford University Press.
- Rolls, E. T. (2008b). Memory, attention, and decision-making, a unifying computational neuroscience approach. Oxford: Oxford University Press.
- Rolls, E. T. (In press). Consciousness, decision-making, and neural computation. In V. Cutsuridis, A. Hussain, J. G. Taylor, D. Polani and N. Tishby (Eds.), *Perception-Reason-Action Cycle, Models, Algorithms and Systems*. Berlin: Springer.
- Rolls, E. T. and Deco, G. (2010). *The noisy brain, stochastic dynamics as a principle of brain function*. Oxford: Oxford University Press.
- Rolls, E. T., Grabenhorst, F. and Deco, G. (2010). Choice and confidence in the brain.
- Rolls, E. T., Grabenhorst, F. and Deco, G. (2010). Decision-making, errors, and confidence in the brain.
- Rolls, E. T., Hornak, J., Wade, D. and McGrath, J. (1994). Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *Journal of Neurology, Neurosurgery and Psychiatry*, 57, 1518–1524.
- Rosenthal, D. M. (1990). A theory of consciousness. (ZIF Report No. 40). Bielefeld, Germany: Zentrum fuer Interdisziplinare Forschung.
- Rosenthal, D. M. (1993). Thinking that one thinks. In M. Davies and G. W. Humphreys (Eds.), *Consciousness* (pp. 197–223). Oxford: Blackwell.
- Rosenthal, D. M. (2004). Varieties of higher-order theory. In R. J. Gennaro (Ed.), *Higher order theories of consciousness* (pp. 19–44). Amsterdam: John Benjamins.
- Rosenthal, D. M. (2005). *Consciousness and mind*. Oxford: Oxford University Press.
- Wang, X. J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*, 36, 955–968.
- Weiskrantz, L. (1997). *Consciousness lost and found*. Oxford: Oxford University Press.

Social consciousness

Kai Vogeley

Experiencing oneself as autonomous agent who interacts with others depends on the cognitive capacities of self-consciousness as awareness of ones own mental states and social consciousness as knowledge of other persons minds. The study of neural correlates of these capacities elicits an impressive overlap with the so-called “default mode of brain function” as a neurobiological universal of mammalian brains. This empirical observation stimulates the speculation that one major cognitive function of the neural default mode is social cognition which conversely implies that humans have a disposition for social cognition that is reflected in this neural default mode.

Mental phenomena, psychological processes and neural mechanisms

Consciousness is a unique phenomenon that can be characterized from a first-person-perspective as the phenomenal experience of one's own mental states including perceptions, judgments, thoughts, intentions to act, feelings or desires. Complementary to this first-person-account of subjective experience, these subjective experiences of mental phenomena can be reconstructed from a third-person-perspective as psychological processes and their underlying neural mechanisms. In other words, whereas mental phenomena are experienced in a subjective space-time-system psychological processes and neural mechanisms are observables in an objective space-time-system.

Following a naturalistic account, which holds that – at least in principle – all relevant cognitive processes and their neural mechanisms can be studied and fully understood by employing natural sciences, consciousness can be conceptualized as being closely linked to cognition thereby bridging the first- and third-person-accounts: Consciousness can be considered “the subjective experience of cognitive function” that “can emerge from the operation of any cognitive function” (Fuster, 2003). Teleologically, consciousness provides an integrated internal representation of the outer world and ones own organism based on experiences and memories providing reflected responses to the needs of our environment (Vogeley et al.,

1999). The search for the underlying neural mechanisms is based on the strong intuition that “conscious experience results from the operation and interaction of several functions in complex assemblies of cortical networks” (Fuster, 2003). Consciousness is constituted by distributed neural processes in different brain regions that might vary from one instantiation of a conscious experience to another: “Any portion of neocortex can generate conscious phenomena as a participant in cognitive function. Thus, the neocortical contribution to consciousness varies from one time to another and from one area to another.” (Fuster, 2003).

Self-Other-differentiation as prerequisite for social consciousness

Survival in a social world clearly requires the ability to distinguish between oneself and others as the “owner” or “agent” of mental states in order not to become confused about one’s own mental states and those of others. This ability is a necessary prerequisite of self-consciousness that can be defined as the (implicit or explicit) awareness of one’s own mental or bodily phenomena as one’s own mental or bodily phenomena. Self-consciousness in this sense includes an immediate self-representation of a cognitive system but this needs not to be explicitly present in the subjective experience. The ability to explicitly meta-represent one’s mental or bodily states as one’s own appears as one special case of self-consciousness.

This human “sense of self” comprises at least the following essential features that can be seen as constituents of human self-consciousness: First, the experience of ownership (with respect to perceptions, judgements etc.) or agency (with respect to actions, thoughts etc.), second, the experience of “perspectivalness” with conscious states being “centered” around myself, and thirdly, the experience of unity forming a long term coherent whole of beliefs and attitudes (Vogeley et al., 1999). The experience of ownership is reflected by the use of a pronominal syntax in language and the experiential quality of agency, that I am performing my movements, having my own perceptions, memories, and thoughts. The experience of ownership (with respect to perceptions etc.) has to be distinguished from the experience of agency (with respect to actions) (Fink et al., 1999). In the case of involuntary action “I may acknowledge ownership of a movement [...] but I may not have a sense of causing or controlling the movement” (Gallagher, 2000). Finally, the experience of perspectivalness refers to the incorporation of my memory, perceptions, and thoughts in my own body and thus to the experience of a literally spatial, body-centered perspective. The experience of unity is associated with long term coherent wholes of beliefs and attitudes that are consistent with preexisting autobiographical contexts, autobiographical memory can serve as empirical indicator here (Fink et al., 1999).

It was postulated that these basic properties are integrated in a so-called “self model” as an episodically active complex neural activation pattern in the human brain, possibly based on an innate and “hard-wired” model. This self model could then plausibly serve as a continuous source of milieu information from the own body and organism, that is activated whenever conscious experiences including properties of ownership, perspectivalness or unity occur (Vogeley et al., 1999).

Intuitive and inferential social consciousness

Being aware of one's own mental states and being able to “read” other persons' minds, to “understand” or “simulate” what they are experiencing are clearly two different capacities and it must be considered how the individual ontogenetic development is related to the social context or to the social origin of self-consciousness. Self-consciousness and social consciousness are functions of a dynamic interplay of social interactions in which the contents of consciousness (of oneself or others) are experienced via automatic attunement processes to others and by comparing and distinguishing own and others' mental phenomena. This leads to the gradual development of self-consciousness (Trevarthen, 2001). This interplay of self-consciousness and social consciousness may then establish a basic form of intentionality that predisposes for social interaction with others and – later in the development – even for the acquisition of detached mental representations of social interactions. This difference can also be conceptualized as primary and secondary intersubjectivity (Gallagher, 2005). Social consciousness thus appears as the fundamental basis of establishing and specifying social relations.

Only a part of psychological processes are experienced as conscious, as shown for instance for the processing of nonverbal communication cues (e.g. Burgoon et al., 1996). One very important distinction that needs to be emphasized is therefore related to different levels of processing of social information that can be either implicit or explicit. Implicit information processing refers to a comparably fast, automatic, prereflexive mode that is employed for instance during non-verbal behaviour. In contrast, explicit information processing comprises processes in a comparably slow, controlled, reflexive, inferential format such as stereotypes or processing based on explicit rules (Lieberman, 2007; Frith & Frith, 2008).

Among nonverbal communication cues gaze behaviour is a salient nonverbal signal for social interest and the intention to communicate, thus indicating mental states of significant others (Argyle and Cook, 1976). “Social gaze” is crucially modulated by various dynamic characteristics among which gaze duration is of particular relevance. In an own study we investigated the neural correlates of participants while being gazed at with varying gaze duration. Increasing gaze

duration resulted in an increased likeability rating and increased neural activation in the medial prefrontal cortex (MPFC) (Kuzmanovic et al., 2009), a region that is well known to be involved in mentalizing and social cognitive tasks. This process of gaze evaluation can be understood as comparably late stage of processing implemented in the MPFC as “inexact, probabilistic, internally generated” cognition (Mitchell, 2009) that crucially depends on the personal history of the experiencing person.

A very prominent strategy to study inferential, rule based ascription of mental states to others are so-called “theory of mind” paradigms that require the modelling of the knowledge, the attitudes or beliefs of another person prospectively, a key brain region involved here is again the MPFC (Amodio and Frith, 2006). However, these levels of processing should be conceptualized as a continuum of different levels of processing resulting in a flexible capacity of person perception that is performed either automatically or in a controlled manner, depending on the data available.

True interaction

Gaze can also establish triadic relations between self, other and the world via the modulation of attention of others or “joint attention”, an important precursor of social cognition (Moore and Dunham, 1995). It has been suggested that it might be the motivation to participate in joint attention that leads to shared, social realities as a unique aspect of human cognition (Tomasello et al., 2005). To realize a truly interactive paradigm during ecologically valid conditions (Singer, 2006) we developed an eyetracking setup allowing to track a participants’ gaze behavior and to contingently control the gaze behavior of a computer-animated character visible on the screen.

Participants were instructed – while undergoing functional neuroimaging – to lead the gaze of the other person towards one of three objects by looking at it. The gaze behavior of the other was made responsive to the participant’s gaze and varied to either follow it (joint attention) or not. Alternatively, participants were asked to respond to the other by looking at the same or at another object. Employing this technique we were able to study the neural correlates of the reciprocal nature of social interaction. As one of two key findings, our analysis of the fMRI data demonstrated that joint attention is associated with a recruitment of the MPFC, the second interesting finding was that directing someone else’s gaze towards an object activated the ventral striatum corresponding to the hedonic and motivational aspects of sharing attention with others. Another interpretation would suggest that this finding reflects that control of others is an important

motivation for engagement with other people and for trying to make sense of them (Fiske and Dépret, 1996). Consequently, our findings may point towards a putative relationship of social motivation resulting from the interpersonal coordination of behavior and social influence as the control one has over the behavior of another person (Schilbach et al., in press).

Neural mechanisms of social consciousness

Cognitive neuroscience of the last decade has revealed that essentially two different systems are recruited during social cognitive processes, namely (i) the “social neural network” (SNN) including the MPFC, the temporoparietal junction and superior temporal sulcus, and the temporal pole (Lieberman, 2007), and (ii) the so-called “mirror neuron system” (MNS) including superior parietal and premotor regions. Whereas the attribution of mental states to others appears to recruit SNN, the MNS shows increased neural activation when motor components are involved as in real or virtual actions or their simulations (Keysers and Gazzola, 2007). MNS activation seems to correlate with “early” stages of social cognition such as the detection of motor expertise and might putatively also underlie the fast processing of “first impressions” on the basis of facial expressions or gestures. In contrast, the SNN is recruited during comparably “late” stages of evaluation of socially relevant information. Social cognition appears to constitute a “natural kind” a cognitive domain that is characterised by “the dependence on a qualitatively distinct class of mental representations” (Mitchell, 2009), seemingly different cognitive phenomena (including thinking about oneself, accessing one’s attitudes, experiencing emotions, inferring other person’s mental states etc.) have in common the characteristics of “fuzzy”, probabilistic and internally generated cognition, on a neural level they share the same substrate, the MPFC (Mitchell, 2009).

Notably, these key regions of the SNN have also demonstrated to be active during so-called resting states or baseline conditions that are characterized by the absence of any external instruction of an experimenter. This specific pattern of distribution of baseline activation has been identified as the so-called “default mode of brain function” or default network (DN) and was first studied by Raichle et al. (2001). Meta-analytical studies (Buckner et al., 2008) demonstrate a significant overlap between SNN and DN. Intriguingly this default network (DN) can be observed in humans irrespective of the task they are involved in and the methodology with which brain activity is measured and is also observed in other mammals (Vincent et al., 2007) and can thus be considered a neurobiologically universal building principle of mammalian brains.

Conclusions

Of course, such a neurobiological universal like the DN is thought provoking and stimulating. On the basis of the striking finding of a considerable overlap of the SNN and DN one might be tempted to speculate that at least one key psychological function of the neurobiological DN is social consciousness. Admittedly, the empirical evidence for the speculation on such an association between the neural mechanisms of social consciousness on the one hand and the default mode of brain function on the other hand is weak at the moment. The key problem that arises is that we do not have access to the phenomenal experience or the psychological processes during resting states. This problem is inherent to resting states: As soon as we would systematically approach the inner experience during resting states, defined by the absence of any sort of external task or instruction, these resting states would be disturbed and would immediately stop to be what they are. This absence of any behavioral measure is in strong contrast to classical functional neuroimaging studies that are performed on the basis of well-controlled psychological experiments in which only on the basis of a clear description of psychological processes the functional role of the neural correlates can be identified. We can consequently – in the absence of adequate research strategies that allow for the systematic study of the phenomenal experiences of the DN – only speculate about social information processing as the inner experiences during resting states. However, based on the observation of an impressive overlap of SNN and DN this speculation does not appear empirically completely unjustified, in fact, the key argument of the vast majority of functional neuroimaging studies in cognitive neuroscience is the topographical distribution of brain regions involved in – usually well-defined – psychological processes.

This convergence supports the hypothesis that we have a disposition for social cognition that is neurobiologically instantiated and potentially disturbed during disturbances of social cognition. An even more ambitious move would be the following: Let us assume that it was true that, first, DN is a neurobiological universal in humans (and in some variants putatively even in other mammals) and that, second, social consciousness is a key function of the DN. This would in conclusion imply that social consciousness itself is a universal cognitive capacity in humans (and in some variants putatively even in other mammals). Most interestingly, this is exactly what has been proposed by Michael Tomasello on the basis of comparative behavioral studies in humans and non-human primates. He proposes that social cognitive capacities as a prerequisite for human “culture” (in a universal sense) are relevant for the whole genus of *homo sapiens* (Tomasello et al., 2005). Social consciousness is what makes us human, and social cognitive neuroscience studies the neurobiological mechanisms that underly these specifically human cognitive capacities.

References

- Amodio, D. M. and Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, 7, 268–277.
- Argyle, J. M. and Cook, M. (1976). *Gaze and Mutual Gaze*. Cambridge: Cambridge University Press.
- Buckner, R. L., Andrews-Hanna, J. R. and Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals New York Academy of Sciences*, 1124, 1–38.
- Burgoon, J. K., Buller, D. B. and Woodall, W. G. (1996). *Nonverbal communication: The unspoken dialogue* (2nd Ed.). New York: McGraw-Hill.
- Fink, G. R., Markowitsch, H. J., Reinkemeier, M., Bruckbauer, T., Kessler, J. and Heiss, W. D. (1996). Cerebral presentation of one's own past: Neural networks involved in autobiographical memory. *Journal of Neuroscience*, 16, 4275–4282.
- Fiske, S. T. and Dépret, E. (1996). Control, interdependence, and power: Understanding social cognition in its social context. In H. M. Stroebe W. (Ed.), *European Review of Social Psychology* (pp. 31–61). New York: Wiley.
- Frith, C. D. and Frith, U. (2008). Implicit and explicit processes in social cognition. *Neuron*, 60, 503–510.
- Fuster, J. M. (2003). *Cortex and mind – Unifying cognition*. Oxford: Oxford University Press.
- Gallagher, S. (2005). *How the body shapes the mind*. Oxford: Oxford University Press.
- Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Science*, 4, 14–21.
- Keysers, C. and Gazzola, V. (2007). Integrating simulation and theory of mind: From self to social cognition, *Trends in Cognitive Science*, 11, 194–196.
- Kuzmanovic, B., Georgescu, A. L., Eickhoff, S. B., Shah, N. J., Bente, G., Fink, G. R. and Vogeley, K. (2009). Duration matters: Dissociating neural correlates of detection and evaluation of social gaze. *Neuroimage*, 46, 1154–1163.
- Lieberman, M. (2007). Social cognitive neuroscience: A review of core processes. *Annual Reviews Psychology*, 58, 259–289.
- Mitchell, J. P. (2009). Social psychology as a natural kind. *Trends in Cognitive Science*, 13, 246–251.
- Moore, C. and Dunham, P. J. (1995). *Joint Attention: Its Origin and Role in Development*. Hillsdale: Lawrence Erlbaum.
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A. and Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences USA*, 98, 676–682.
- Schilbach, L., Wilms, M., Eickhoff, S. B., Romanzetti, S., Tepest, R., Bente, G., Shah, N. J., Fink, G. R. and Vogeley, K. (in press). Minds made for sharing: Initiating joint attention recruits reward-related neurocircuitry. *Journal of Cognitive Neuroscience*.
- Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research. *Neuroscience and Biobehavioral Reviews*, 30, 855–863.
- Tomasello, M., Carpenter, M., Call, J., Behne, T. and Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28, 675–735.
- Trevarthen, C. (2001). The neurobiology of early communication: Intersubjective regulations in human brain development. In: A. F. Kalverboer and A. Gramsbergen (Eds.), *Handbook on Brain and Behavior in Human Development*. Dordrecht, The Netherlands: Kluwer.

- Vincent, J. L., Patel, G. H., Fox, M. D., Snyder, A. Z., Baker, J. T., Van Essen, D. C., Zempel, J. M., Snyder, L. H., Corbetta, M. and Raichle, M. E. (2007). Intrinsic functional architecture in the anaesthetized monkey brain. *Nature*, 447, 83–86.
- Vogeley, K., Kurthen, M., Falkai, P. and Maier, W. (1999). The prefrontal cortex generates the basic constituents of the self. *Consciousness and Cognition*, 8, 343–363.
- Vogeley, K. and Roepstorff, A. (2009). Contextualising Culture and Social Cognition. *Trends in Cognitive Science*, 13, 511–516.

Consciousness and language

A processing perspective

Michael Sharwood Smith and John Truscott

Our conscious control of language and its acquisition is strictly limited. A processing-oriented perspective to explain this will be outlined called MOGUL according to which some linguistic processes are inherently unconscious while others can be either conscious or not. The former involve representations, found in the dedicated, uniquely human language module, whose sealed-off nature does not permit the activation levels necessary for conscious experience to take place. However, language knowledge created by cognitive processes that are not specifically linguistic, that is, created outside the language module, can indeed be raised to consciousness. This is because they are more directly connected to the perceptual system and are accordingly open to much higher levels of activation.

It is easy to take language for granted. When listening to speech, all that 'goes into our heads' is a sequence of minute changes in air pressure moving the basilar membranes in our ears. By the same token, as the reader reads these words, no words are going in but rather subtle changes in light waves are being picked up in the retinas. Yet, as humans, we can decode these subtle events, recreating intended messages from mere puffs of air and patterns of light. Language represents a highly complex and peculiarly human ability. Apart from using it as a means of communicating and organising our thoughts, we are also able to reflect on language as an object and talk about it so that consciousness and language seem to go hand in hand. Nevertheless, the evidence strongly suggests that our conscious control of language and its acquisition is strictly limited. We will briefly review the way linguists have addressed the issue of consciousness. A new perspective will then be outlined, reconciling two views on consciousness, one developed by Bernard Baars and one by the linguist, Ray Jackendoff (Baars, 1988; 1997; Jackendoff, 1987; 2007). The new perspective is called MOGUL.

The biological approach to language

The most dominant force in modern linguistics has been, and arguably still is generative grammar (Newmeyer, 1996). Generative linguists conceive of language as a psychological and biological phenomenon. The main goal of linguistic theory as Chomsky defined it is to explain the principles that constrain the form that human languages can take such that cognitively immature children, with no ability or desire to consciously analyse what they are doing, can acquire complex grammatical systems without any need for correction by adults (Chomsky, 1975). Both humans and many animals share an ability to link sounds with meaning but it seems that humans, unlike other species, have evolved a special linguistic ability that is sufficiently important for their survival to develop very early in life. Hence, in this view, the goal of linguistics is to explain the properties of this biologically endowed system (Anderson and Lightfoot, 1999). Although the neural correlates of this language faculty are not fully understood, certain areas of the brain have long been associated with various language functions, in particular Broca's and Wernicke's areas, and there is a fast growing area of research using such techniques as event related potentials and functional magnetic resonance imaging to try and associate specific linguistic functions with various types of brain activity (e.g. Kutas et al., 1999; Mueller, 2005).

How to know a language and at the same time, not know it at all

The generative view of language implies that knowledge, use, and acquisition of grammar is largely subconscious (see Roberts, 1994). Perhaps for this reason, issues of consciousness have received little attention in this area. The primary exception is the work of Ray Jackendoff (1987 and 2007), which we will return to shortly. Regarding second language acquisition, consciousness has attracted some attention. The fact that older learners engage in more conscious reflection when learning a language and yet appear to lack the linguistic ability of toddlers has led to the claim that older learners have lost the 'language instinct' (Pinker, 1994) and therefore must resort to less effective, conscious ways of mastering the new system (Selinker, 1972). This claim has been challenged by Stephen Krashen (Krashen, 1981) pointing to evidence of fixed developmental patterns which child language researchers had noted also turning up in the spontaneous, as opposed to the slow conscious reflective language performance of adult second language learners (Dulay et al., 1982). This led to Krashen's claim that there were two entirely separate developmental processes at work, a *conscious* one which allowed learners to use consciously learned grammatical rules to edit and correct their

performance albeit in very limited, laborious ways and a *subconscious* one which drove spontaneous language performance.

The subconscious processes involved are those that also drive first language acquisition and have, by hypothesis, survived into later life. Thus adults' impression that they are promoting their development by paying conscious attention to grammar is essentially an illusion. You may be perfectly aware of what the correct word order for a German sentence should be but still, frustratingly, go on making the same old errors. Conscious knowledge may allow you to spot and correct your mistakes, if you have time, but it does not stop you making them again later. Rather, grammatical development follows its own course often proceeding through various stages having nothing to do with learners' or teachers' intentions. Second language acquisition research within a Chomskyan perspective has assumed the necessarily subconscious acquisition of grammar as a fact of life for both younger and older learners. However, others have adopted models taken from cognitive psychology which deny language a special status in the mind/brain. Many claim that conscious learning of grammar can, with practice, be transformed into subconscious, automatised routines in the same way we appear to learn how to drive a car or any other skilled activity (Johnson, 1996). But for those in second language acquisition research who subscribe to the existence of a special language faculty, the question of whether and how much the language instinct survives in later life is a live issue. There are numerous positions on this issue but the one taken in this chapter is, in broad agreement with Krashen's claim, that linguistic ability operates at two distinct levels. One is handled by a closed-off language module, the operations of which are inaccessible to consciousness. It develops during our lifetime, remains with us for acquisition of new languages and controls much of the way we produce and understand language. Then there are numerous aspects of language outside the language module itself which can give rise to conscious experience in ways that will be discussed below.

MOGUL

Our proposal is based on the architecture of the language module as conceived by Jackendoff and is developed within a research framework called MOGUL (Modular Online Growth and Use of Language) (Truscott and Sharwood Smith, 2004). The module has two main subcomponents, each modular in itself in the sense that they operate separately and blindly (hence fast and efficiently) with their own particular code. One is concerned with the organisation of speech sounds (phonology) while the other dictates permissible ways of combining and sequencing word forms to create sentence structure (morphosyntax). Interface processors

- (a) match up the output of the two linguistic subsystems with each other and
- (b) match up these linguistic outputs with elements outside the language system, i.e. thoughts at the higher level and peripheral systems to do with speech, signing and writing.

Let us begin with the way thought is converted into sound. Meanings, here termed *conceptual structures* (CS), are matched up with elements of morphosyntax, written in a code called *syntactic structure* (SS). The matching up or 'co-indexing' of CS with SS is done by an interface processor that associates conceptual structures with syntactic structures. For example, in the case of *the dog*, the conceptual structure underlying the meaning of 'a certain specified dog' is matched up with article plus noun. In this way a processing chain begins to be formed. Another interface processor coindexes this syntactic structure and elements of the other linguistic code, *phonological structure* (PS), yielding a chain $CS \leftrightarrow SS \leftrightarrow PS$. Take the sentence, *the dog needs a walk*. The conceptual structure underlying the idea that the dog needs a walk is matched up with a string of syntactic elements, expressed in traditional terms as Article + Noun + Verb + Article + Noun, the precise syntactic structure being considerably more complex than this. This syntactic string is then matched up by the SS/PS interface with elements of phonological structure, informally expressed as /ðə dɒg ni:dz ə wɔ:k/. After this, the specifically linguistic part of the processing is finished. The phonological structure is then linked up to the articulatory systems responsible for speech or writing. Comprehension is a process going in the opposite direction, using the same system. In both cases, as the structural chain representing the message is built up incrementally there is much to-ing and fro-ing as some matches are discarded for ones which fit better. All of this is inaccessible to consciousness and cannot be sidetracked or stopped except where no match can be made.

Each subsystem consists of a processor and an information store which can be thought of as a specialised memory. Hence the phonological module has a phonological processor that deals exclusively with items in the phonological store (memory), i.e. phonological structures (PS). Some of these structures are basic building blocks for PS and supplied in advance, i.e. innately. Others are combinations of these elements which develop as a result of experience. A Farsi speaker will therefore have a different PS store than a French speaker: although the basic building blocks are the same, they are combined in different ways. The same holds for SS. A syntactic processor handles and manipulates only elements in its syntactic memory store.

Three important additional points must be made. Firstly, the distinction between working memory and long term memory is captured by defining working memory as a buffer or 'blackboard' containing currently elements in a given memory that have been activated. The second point concerns levels of activation.

Well-established elements in memory will have a high *resting level* and will be easily accessible for processing; less well established, i.e., new or infrequently used elements will typically have lower resting levels. Finally, in a speaker's or listener's attempts to process language, there is constant competition between different elements for selection in working memory. In this race, those elements with a higher resting level win out against those that are less accessible. However, repeated use of a PS or SS will result in a gradual increase in its resting level and hence its relative accessibility. This last feature is captured by Acquisition by Processing Theory (APT), a key part of the MOGUL approach (Truscott and Sharwood Smith, 2004).

Thus far we have focussed on the language module and its connections to CS. Another part of MOGUL architecture crucial for discussion of consciousness is *perceptual output structures* (POpS). It is a highly interconnected set of perceptual blackboards, each representing the ultimate output of a sensory system. *Auditory structures* (AS), for example, constitute the output of the system that deals with acoustic events in the environment. AS may be thought of as 'auditory memory' and also underlies sounds experienced in dreams and hallucinations. The rich interconnections in POpS produce a tendency toward synchronization of the component stores, with a resulting focus on a single sensory representation at a time. This focus in turn results in exceptionally high current activation levels for that representation, a point that will become crucial in the following section.

Conscious and unconscious processes in MOGUL

Strictly speaking, there is no such thing as a conscious process, as the objects of consciousness are representations, not processes. The question, then, is which processes involve conscious representations, which do not, and how the two interact. Regarding the 'which' question, MOGUL adopts perhaps the simplest possible position: A representation is conscious when its current activation level is above a threshold. As noted above, the place where such levels can occur is POpS, because the various perceptual stores tend to synchronize their processing on a single representation at a time. This representation can then be used by a variety of processors. This account correctly predicts that conscious experience centers around perception – the *perceptual bias* of consciousness (Baars, 1988). With these characteristics, POpS roughly fills the role of Baars' global workspace (GW) – a central facility with different cognitive processors or coalitions of processors competing to gain access to it. Information that gains access to the GW is broadcast to a range of other specialised processors enabling the mind/brain to rapidly collate information from a range of different sources. GW, in this view, is

the source of conscious experience. Our account also accommodates Jackendoff's 'intermediate level theory of consciousness', according to which consciousness occurs at a level between peripheral, sensory processes and deeper cognitive processes associated with the prefrontal cortex (Jackendoff, 1987). This intermediate level can reasonably be equated with POpS.

The footprints of language

Our dominant conscious experience of language is 'the voice in the head', also experienced in thinking, dreaming and hallucinating. Listening to the voice, we hear individual sounds, stress and rhythm, and changes in pitch, and in this sense we are conscious of language. But we have no inkling of the way our peripheral processing mechanisms have handled the speech signal. Nor do we have any awareness of creating complex syntactic and semantic representations or of how the thoughts that follow on from this have mysteriously arisen in our heads. We have only the voice, the footprints of language, so to speak.

Where does this voice come from? Its source cannot be the language module itself, as the module has nothing like the conditions that allow POpS representations to reach the extreme activation levels required for consciousness. Processes using PS and SS representations are inherently unconscious. This conclusion is supported by the fact, noted by Jackendoff, that we can be aware only of certain features of speech sound; the fine structural detail found in a PS representation is inaccessible to awareness. The natural explanation is that the voice is a product of auditory structures, outside the language module. AS representations, by definition, express sounds but do not incorporate the features of the systems (phonological or conceptual, for example) that utilize those sounds. The voice, then, is a series of AS representations, the perceptual footprints of linguistic processing.

Given that PS and SS representations cannot attain the activation levels necessary for consciousness, it is entirely reasonable that, when we comprehend or produce utterances, we have no notion of the sophisticated mechanisms at work. Nor is it surprising that children, with minimal conscious analytic abilities, can acquire these sophisticated mechanisms. But if the linguistic processes are not themselves conscious, conscious representations in POpS are still important for their operation. In language comprehension, auditory processing yields an auditory (AS) representation, which then serves as input for linguistic processing, that is, in PS and beyond. When the AS does not become conscious, it does not fill this role effectively, because in order for the language module to fully use the information that a POpS representation contains, that representation must be strongly present for a reasonable period. A weak, fleeting appearance might trigger some

limited activity in the language module but would allow little opportunity for processors to deal with the information it contains. This is *subliminal perception*: some information gets through, to be ultimately represented at CS, but it is quite limited. In sum, processing within the language module is entirely unconscious but nevertheless relies on conscious perceptual processes to provide its input and leaves conscious footprints in the form of the voice in the head.

The situation is similar, but interestingly different, when linguistic knowledge outside the language module is involved. Such knowledge consists primarily of conceptual (CS) representations. They are necessarily unconscious, so this means that processes involving extramodular linguistic knowledge can be conscious only to the extent that they use perceptual (POpS) representations and do so in a way that involves high activation levels. Suppose, for example, that someone with limited Spanish ability has learned the rule that Spanish adjectives follow nouns. The CS representation of this information is not itself conscious, but it can be given linguistic form, including an auditory (AS) representation such as [*adjective follows noun*], coindexed with it. This perceptual representation can become conscious. The conscious experience, the 'voice in the head', corresponds to activation of the inherently unconscious conceptual representation. Extramodular linguistic knowledge and its use thus involve both conscious and unconscious processes, interacting. At the same time, it cannot, by definition, be merged with the separate system developing within the language module itself. It also remains an open question exactly how much extramodular knowledge of language can enhance or hinder performance.

The conscious experience of the rule's use can take various forms. Suppose, for instance, the person wants to say *red book* in Spanish and knows the words *rojo* (red) and *libro* (book). The initial conscious experience might involve the sound of the two Spanish words (AS representations), possibly preceded by that of the corresponding English words. These might be accompanied by a visual image of the written forms of the words (VS representations). If this image reflects the English order, as is likely for a low-level Spanish learner, it could be followed by a visual perception of movement, something like dragging and dropping, yielding a visual (and possibly auditory) experience of *libro* and *rojo* in their proper order. The movement might be accompanied by feelings from muscles associated with eye movement (whether or not such movement actually occurs). The details no doubt vary greatly, and though these events are slow in neural terms, their speed in experiential terms often makes it difficult to pin down exactly what is occurring.

The conceptual (CS) representation of the rule, the key part of these experiences, is never directly present in the conscious experience. This situation is in fact typical of ordinary thought, which involves an interplay of conscious and

unconscious processes, as seen in the way the voice in the head figures in thinking. The conscious experience (the voice) is of AS representations. Because they are coindexed with CS representations, their activation correlates with activation of the latter. Thus the conscious experience of the perceptual representation is intimately associated with the unconscious conceptual representations. The conceptual information is not conscious but is strongly present in the background.

Returning to the case of Spanish word order, whatever perceptual experiences occur tend to fade with practice. Initial application of the rule is a conscious process, but with practice it becomes steadily quicker and less effortful, and these changes are accompanied by a decline in awareness. This is *automatization*, which in MOGUL terms is a gradual increase in the resting activation levels of the relevant CS representations, making them more readily available to processors. They enter into processing more quickly and with less stimulation from POpS representations and so can be used before the current activation levels of the latter have risen to the level of consciousness (a slow process, in neural terms). Automatization is one of the *redundancy effects* that Baars (1988) offered as evidence against a pure activation account of consciousness, the logic being that, in automatization, awareness fades as activation levels presumably rise with practice. In our account this is not a problem because the representations that rise in activation level are in CS while those that are potentially conscious are in POpS.

Extramodular linguistic processes typically begin as conscious, because the development of conceptual knowledge depends on relatively strong, enduring perceptual representations, which are for this reason conscious. When they are not conscious, or only fleetingly conscious, the development of new knowledge based on them is quite limited. This is *implicit learning*. It occurs when circumstances do not allow the relevant perceptual representations to reach the level of activation needed for awareness. Such circumstances are deliberately created in laboratory experiments in which either the rule is too subtle to be consciously inferred or the person's attention is focused elsewhere. Such designs prevent the POpS representation from achieving a sufficient activation level. The result is that any new CS representation is likely to be poorly developed and to have only a minimal resting activation level. Such a representation can influence processing by raising the activation levels of related representations and thereby increasing the likelihood of their use (priming), but it cannot be directly associated with a conscious experience because of its minimal activation level, greatly limiting its role in processing.

Conclusion

Some linguistic processes are inherently unconscious while others can be either conscious or not, depending on the circumstances. The former involve representations, found in the language module, that can never reach the necessary activation levels. We are aware only of their footprints, i.e. the voice in the head. Modular linguistic knowledge lies at the heart of human linguistic ability. Extramodular linguistic knowledge, taking the form of conceptual (CS) representations, is also unconscious in itself, but its close association with perceptual representations gives it intimate involvement in conscious experience. If this extramodular knowledge has been used extensively, its continuing use may not result in associated perceptual representations achieving the activation needed for awareness, weakening or eliminating the connection with consciousness. Use of these representations can therefore be either conscious or unconscious. Finally, conscious and unconscious processes interact in two major ways: In language comprehension, conscious perceptual processes provide the raw material for unconscious processing in the language module, and the use of extramodular knowledge involves a constant interplay of unconscious conceptual processes and (variably) conscious perceptual processes.

References

- Anderson, S. R. and Lightfoot, D. W. (1999). The human language faculty as an organ. *Annual Review of Physiology*, 62, 697–722.
- Baars, B. (1988). *A cognitive theory of consciousness*. New York: Cambridge University Press.
- Baars, B. (1997). In *the theater of consciousness: The workspace of the mind*. New York: Oxford University Press.
- Chomsky, N. (1975). *Reflections on language*. New York: Pantheon.
- Dulay, H., Burt, M. and Krashen, S. (1982). *Language two*. New York: Oxford University Press.
- Jackendoff, R. (1987). *Consciousness and the computational mind*. Cambridge, MA: MIT Press.
- Jackendoff, R. (2002). *Foundations of language*. Oxford: Oxford University Press.
- Jackendoff, R. (2007). *Language, consciousness, culture: Essays on mental structure*. Cambridge, MA: MIT Press.
- Johnson, K. (1996). *Language teaching and skill learning*. Oxford: Blackwell.
- Krashen, S. D. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.
- Kutas, M., Federmeier, K. D. and Sereno, M. I. (1999). Current approaches to mapping language in electromagnetic space. In C. M. Brown and P. Hagoort (Eds.), *The neurocognition of language* (pp. 359–392). Oxford: Oxford University Press.
- Mueller, J. (2005). Electro-physiological correlates of second language processing, *Second Language Research*, 21, 152–174.

- Newmeyer, F. (1996). *Generative linguistics: A historical perspective*. London: Routledge.
- Pinker, S. (1994). *The language instinct: The new science of language and mind*. London: Penguin.
- Roberts, I. (1994). Universal Grammar and L1 acquisition. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 455–475). London: Academic Press.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, 10, 209–231.
- Truscott, J. and Sharwood Smith, M. (2004). Acquisition by processing: A modular approach to language development. *Bilingualism: Language and Cognition* 7, 1–20.

Cognitive illusions

From magic to science

Gustav Kuhn

Magicians have developed powerful techniques to manipulate our perception and awareness. Many of these techniques share similarities with phenomena typically investigated by psychologists and neuroscientists. Here a novel approach to the study of consciousness is proposed which utilizes the magician's expertise to complement more traditional experimental laboratory based research. It is argued that this approach offers new and exciting insights into wide areas of consciousness, such as attention, visual awareness and top-down modulation of perception.

Throughout history different theoretical and empirical approaches have been used to study the human mind. For example, patients with localized brain damage may exhibit specific cognitive impairments, which provide insights into the cognitive function of the normal brain. The main interest here is not the lesion itself, but what it can tell us about the "normal" mind. In contrast, rather than investigating impairments, we may study individuals with particular expertise. For example, it has been shown that people with domain specific expertise encode and represent domain specific information differently thus suggesting that expertise results in altered perceptual experience, which give us clues on how we learn to see (e.g. Chase and Simon, 1973). Perhaps the strongest and most powerful examples of distortions in perception are those produced by magicians. Magicians have developed ways to distort our perception often resulting in wonder and disbelief (Parris et al., 2009). Indeed distortions of awareness are arguably the core definition of magic. Given this close link between magic and awareness, it is rather surprising that with few exceptions (e.g. Kuhn and Land, 2006; Kuhn and Tatler, 2005; Triplett, 1900) this relationship has been largely ignored by modern science. However, as I will argue here, and have elsewhere (Kuhn et al., 2008), much can be learnt from the way in which magicians distort our perception. In this paper I will demonstrate how magic and science can be

combined to form a science of magic. In particular I will argue that this rather novel approach promises new and exciting avenues to discovering more about human consciousness.

Magicians are not unique in manipulating our awareness. Several “experts” have learnt to play tricks with our mind. For example, film directors have developed wonderful techniques to distort our visual perception, many of which share similarities with those used by magicians (Kinlsey, 1993). Pickpockets manipulate their victim’s tactile awareness, using techniques that parallel those of the conjuror. Advertising agents manipulate our decisions, whilst first-rate sports men/women frequently deceive their opponents thorough deceptive movements (e.g. Jackson et al., 2006). The only difference between the conjuror and these other experts is that magicians have made distortions of awareness the centre of their art. They have experimented with these principles for several millennia (Christopher, 2006) and have built up a solid tradition of distorting reality. The overlap between the conjurors’ principles of deception and those used by others demonstrates that these techniques don’t merely apply to the limited context of the magic performance, but are general principles of cognition.

The first step towards a science of magic requires us to identify the link between the magic trick and the scientific principles involved. This process requires translating the conjuror’s language into that of the scientist (Kuhn et al., 2008; Macknik et al., 2008). For example, conjurors frequently use misdirection to prevent the audience from detecting their secrets. They define misdirection as a “deflection of attention for the purpose of disguise” (Sharpe, 1988). Visual scientists on the other hand would describe misdirection as “attentional modulation of visual awareness”. Whilst the languages may differ, the underlying principles are the same. Rather than processing all of the perceptual information, we only select the information that is of importance, through systematic orienting of attention. Whilst this attentional selection process is a very efficient way of dealing with vast amounts of sensory information, it does come at a cost; namely unless we attend to something, information about it is rapidly lost. This point is vividly illustrated by the change blindness phenomenon, whereby people often fail to see visually salient changes to a scene when the change occurs with a visual disruption of the scene. This demonstrates that whilst our conscious experience of the world is one of rich sensory detail, our conscious representation is far less complete (Rensink et al., 1997). What we see is therefore largely determined by what we attend to, and this attentional orienting can either take the form of overt eye movements (overt attention), or allocating additional processing resources in the absence of eye movements, known as covert attention. Misdirection may therefore provide insights into areas such as attentional orienting and awareness, which lie at the heart of the study of consciousness.

Once this link between magic and science has been established, we can take advantage of the magician's real world experience to investigate questions that are of scientific interest. Many conjuring techniques, such as misdirection, are complex multi-layered principles, which often makes them rather difficult to investigate scientifically. The second challenge involves finding ways of breaking down some of these techniques so that they can be studied in a scientifically meaningful way. Whilst observations and descriptions are important starting points, the true value of magic is only unleashed by incorporating conjuring techniques into the scientific paradigm. The approach we propose involves identifying the principles used by magicians and then using a more reductionist approach to investigate it scientifically so as to learn more about the underlying mechanism. These findings can then be used to guide the more traditional experimental research, thus closing the gap between magic and science, and thus forming a science of magic.

Over the past five years we have developed several paradigms in which conjuring techniques are employed to manipulate people's visual perception. I will talk about two of these, namely misdirection and illusions.

Misdirection

Misdirection is a complex multilayered process in which people's attention is systematically orchestrated using a wide range of techniques. In order to study misdirection scientifically we have developed a very simple "magic trick" in which misdirection is used to prevent people from seeing a fully visible event (Kuhn and Findlay, 2009; Kuhn and Tatler, 2005; Kuhn et al., 2009; Kuhn et al., 2008). In one of these tricks a magician is seen making a lighter disappear, an effect that was achieved by distracting people's attention at the moment in which the lighter is dropped into the magician's lap (Figure 1). The simplicity of this trick allows us to study misdirection in a scientifically meaningful way, by both identifying the factors that contribute to its effectiveness, as well as manipulating it experimentally.

In order to study something scientifically, we must be able to measure its effectiveness. The most obvious measure of "effectiveness" is whether it succeeds in preventing people from seeing a "to-be-concealed" event. In these paradigms, we ensured that the "to-be-concealed" event was fully visible, which means that participants' detection of the event (i.e. whether they saw the object drop) provides an index of its effectiveness. Moreover, by assuming that detection of the drop requires covert attention (attention without eye movements), we can use participants' detection of the event as an index of covert attention (Tse, 2004). Whilst this awareness probe informs us about the effectiveness of the attentional manipulation at one particular point in time, we can use eye tracking techniques

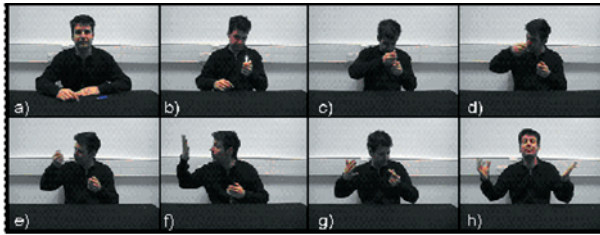


Figure 1 shows the sequence of events in the misdirection trick (Kuhn and Findlay, 2009). The effect in this magic trick was the disappearance of a lighter. In this trick the magician (a) picks up a lighter with his left hand and lights it (b). He then pretends to take the flame with his right hand (c), and gradually moves it away from the hand that is holding the lighter (d, e). During this maneuver the magician is looking at the right hand. Once it has reached the other side, he snaps his fingers, waves his hand and reveals that it is empty (f). At the same time the lighter is dropped into the lap. The drop takes place in full view. The magician now directs his gaze to his left hand (g), raises it, and snaps his fingers to reveal that his left hand is now also empty and the lighter has disappeared (h).

to measure where people are looking through the entire presentation of the trick. The eye tracking data therefore provides us with an online measure of where participants are attending overtly.

The misdirection paradigm allows us to study aspects of the magician's misdirection in a naturalistic, yet scientifically controlled way. It therefore provides us with a valuable tool to investigate various aspects of attention and awareness from a rather novel angle. For example, we have used this paradigm in "real world" situations in which participants faced a real person performing the misdirection trick live (Kuhn and Tatler, 2005). Most of the participants who watched the misdirection trick for the first time failed to see the dropping object even though the event took place right in front of their eyes. The misdirection employed by the magician was, therefore, effective at preventing people from seeing the event. Importantly, participants did not miss the event due to low visibility. For one, when the trick was repeated, most of the participants immediately saw the item being dropped, thus demonstrating that it was indeed visible. Moreover, eye movement records revealed that participants' ability to detect the drop was not related to where they were looking at the moment of the drop, and therefore is not determined by the low level visual input received by the visual system. Misdirection is a powerful demonstration of "inattention blindness", in which people who are engaged in attentionally demanding distractor tasks frequently fail to see unexpected yet fully visible events (Mack and Rock, 1998). However, instead of drawing attentional

resources away through the use of distractor task, misdirection works by implicitly orienting attention away from the to-be-concealed event.

In later work we used pre-recorded versions of similar misdirection tricks which allowed us to control certain aspects of the stimuli more tightly. Whilst these pre-recorded versions of the trick were still effective at misdirecting people's attention, its effectiveness was somewhat reduced, thus demonstrating that the engagement in a one-to-one interaction plays an important part in attentional orientation (Kuhn et al., 2008). However, many of the participants still failed to see the visibly dropping object even though it happened in front of their eyes. Rather surprisingly, again detection of the event was independent of where participants were looking thus demonstrating a striking dissociation between covert and overt attention. Whilst it is clear that under experimental conditions overt and covert attention can be dissociated, it is generally thought that under free viewing the main purpose of covert attention is to plan the next saccade, and does not seem to serve any additional role (Findlay, 2005). However, the current results demonstrate that covert attention is instrumental in determining what we see or miss. It illustrates how what we see is not necessarily determined by what we look at, but by what we attend to covertly, and thus demonstrates the important role that covert attention plays in visual awareness. The effectiveness of the misdirection and the high consistency in eye movements also demonstrate that misdirection is very effective at orienting peoples' overt and covert attention.

Given these strong and reliable effects we can then move to ask which aspect of the misdirection is driving this attentional orienting process. For example, we have investigated the role that social cues (i.e. where the magician was looking) have in misdirecting attention (Kuhn et al., 2009). By experimentally manipulating the misdirection trick, we demonstrated that gaze cues are very important at misdirecting attention. Moreover, as the effects found in this paradigm were rather larger than those typically observed in more standard laboratory attention paradigms (e.g. Friesen and Kingstone, 1998), it suggests that we may be underestimating the importance that social cues play in the real world. This real world approach to attention research also allows us to identify features or events that are important in orienting attention that are typically ignored by mainstream science. Thus far our empirical work on misdirection has predominantly focused on systematic orchestration of spatial attention, and thus relatively "low" level sensory information. However, the magician's concept of misdirection also relates to manipulating attention directed towards thoughts, and thus involves higher levels of awareness. By identifying some of these principles, and developing paradigms analogous to the misdirection trick we may learn much more about the mechanisms of this type of misdirection and gain new insights into consciousness.

Illusions

Our conscious perception of the world requires a transformation of the light that falls onto the retina into recognizable objects. Our visual experience is largely determined by the way in which these signals are interpreted by the brain (Gregory, 1968). However, there are certain situations that present specific difficulties in interpreting this information which then give rise to systematic errors. Most of the time, these errors are unnoticed. However, when we do become aware of them they are experienced as visual illusions. Systematic errors in perception can serve as clues as to how the brain generally solves perceptual problems. For example much of what we have learnt about depth perception comes from investigating our susceptibility towards visual illusions (Gregory, 1968). Illusions lie at the heart of magic. Many of the illusions employed by magicians do not necessarily differ from the tricks the mind plays on us in everyday life. What makes them special are the theatrical ways in which they are presented. Moreover, magicians have learnt to maximize their effectiveness by stage-managing the context in which they are viewed. The systematic investigation into these illusions and particularly the way in which conjurors enhance their effectiveness could therefore provide valuable insights into how the visual system solves some of these perceptual problems.

The illusions typically performed by magicians are often rather complicated. The key to studying magic illusions scientifically involves finding illusions that are both simple yet effective, and to discover ways of measuring their effectiveness. An illusion that ticks all of the above boxes is the vanishing ball illusion (Kuhn and Land, 2006; Triplett, 1900). In this illusion a magician is seen throwing a ball up and down a couple of times. On the final throw, he merely pretends to throw the ball when instead it remains secretly concealed in his hand (Figure 2). Rather surprisingly, people often claim to see a ball leave the magicians hand, even though it remained in the hand. This illusion is simple (i.e. merely involves someone pretending to throw a ball in the air), but very effective. Moreover, participants' experience of the illusion (i.e. whether they saw the imaginary ball or not) can be used as a simple measure of its effectiveness.

Using this vanishing ball illusion we have shown that even though the ball did not leave the hand, over half of the participants were convinced they had seen it move towards the top of the screen (Kuhn and Land, 2006). This finding demonstrates how people's perception of an event is based on the expected outcome, rather than the low level sensory input. Perceptual processes that are based on expectations or prediction have their advantage, as they allow for rapid action, which is particularly useful when fast responses are required. However, as demonstrated by the vanishing ball illusion, reliance on such strategies can lead to errors, and

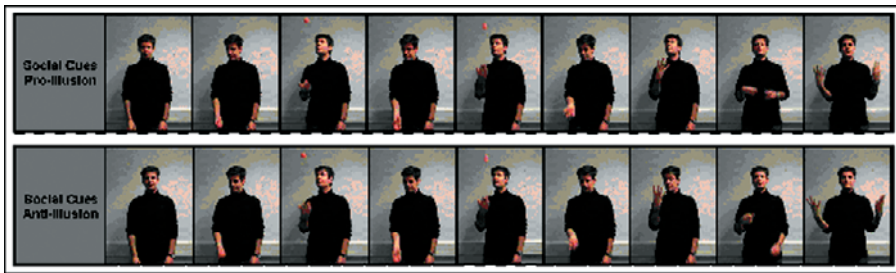


Figure 2 shows the sequence of events in the vanishing ball illusion. In the Social Cues Pro-illusion condition (top panel) the magician's gaze followed the trajectory of the imaginary ball (final pretend throw). In Social Cues Anti-Illusion condition the magician looked at his hand, rather than the illusory trajectory. Figure taken from Kuhn et al. (2008).

illusory percepts. Whilst illusory percepts during a magic show can be rather fun, their effects could have detrimental consequences in other situations such as driving. Thus understanding how we distinguish between illusion and reality may have rather important theoretical, as well as practical, consequences. If our conscious perception is based on expectations it is important to understand the processes that drive these expectations. Conjuring illusions often rely on manipulating expectations, and magicians have learnt to identify situations in which our perception is particularly sensitive to anticipation, and developed ways of manipulating these expectations. The magic literature tells us that social cues (i.e. where the magician is looking) are particularly important in manipulating people's attention as well as intention (i.e. expected outcome) (Sharpe, 1988). Indeed Kuhn and Land (2006) showed that if the magician looked at his hand, rather than the trajectory of the imaginary ball, the illusion's effectiveness is greatly reduced, implying that the social cues did indeed influence expectations, and therefore what was perceived. Moreover, unbeknown to the observers, the eye movement records revealed that participants strategically utilized the social cues to predict the final location of the ball. Rather than merely keeping an eye on the ball, participants glanced at the face prior to fixating on the ball once in the air. The use of adaptive heuristics to solve complex perceptual problems has also been reported in other everyday domains, such as driving (Land and Lee, 1994) or cricket (Land and McLeod, 2000). It is indeed thought that vision in general involves a complex set of problem solving strategies, and the way in which magicians trick these processes may provide important insights into how these strategies work.

Conclusion

Magicians have developed an astonishing expertise in manipulating our awareness. As scientist interested in consciousness and awareness, it would be a shame to ignore this enormous expertise. Luckily the first steps towards recognizing the link between science and magic, as well as the development of paradigms that allow us to test these principles scientifically have been made (Kuhn et al., 2008; Lamont and Wiseman, 1999; Macknik et al., 2008). Whilst the science of magic is indeed a rather novel approach, the concept of using real world expertise to study aspects of the mind is far from new. For example, artists' experimentation with perceptual effects has taught us much about the perception of colour and perspective (Pinna, 2008). Similarly magicians may enlighten us about aspects of attention and awareness. Conjurers are in no way unique, in that they have expertise that could be of scientific interest. However, what makes them special is the dedication that they have given to understand and perfect these perceptual distortions.

Our aim as cognitive scientists is to explain and predict human behaviour in the real world. In cognitive psychology as well as in neuroscience, the trend has been to simplify the issue of investigation by making the experimental context both impoverished and controlled, so as to discover the causal relationships between one factor and the other. The idea that we can learn things about the mind by isolating it in a box and bombarding it with artificial stimuli has however been challenged (Kingstone et al., 2008). In particular, the cognitive ethology framework suggests "that it is by starting with real-world observations and individual variation that one is most likely to generate subsequent research questions for investigation that may lead to general principles of cognition that have relevance to naturally occurring phenomena" (Kingstone et al., 2008). According to this approach, "the initial job of the researcher is simply to observe, describe, and measure what people do and experience in the situation of interest" (Kingstone et al., 2008). In many ways this "job" of observing, describing and measuring human behaviour is something magicians have already done. By utilizing the experience developed by magicians we may, therefore, shortcut this initial, often difficult task of real world observation.

The science of magic as an approach to study consciousness and the human mind in general has several advantages. For one the principles used by magicians are incredibly robust and effective, which means that they are an ideal starting point for scientific explorations. Moreover, many of the techniques and theories are well documented which makes them an ideal starting point for scientific exploration (e.g. Sharpe, 1988). Most of the principles used by magicians have been tried and tested in front of live audiences, which means they are likely to hold

up in the real world, which has major practical applications. Finally, this field of research is still very much in its infancy, but by identifying and investigating some of the techniques used by magicians that may not necessarily be known to the scientific community, this approach has the potential to result in discoveries of new principles.

Much of what we have learnt about consciousness comes from identifying the types of tricks that the mind plays on us. For example, our susceptibility towards change blindness, and inattentional blindness has taught us about the limited conscious visual representation. Similarly investigations into visual illusions such as the hollow mask illusion are powerful demonstrations of how top down information influences our conscious perception. The ease with which magicians can distort our perception is a stark reminder of just how subjective our conscious perception truly is. However, more importantly, by developing a science of magic, which investigates these distortions scientifically, we may gain novel and important insights into some of the factors that are responsible for this conscious representation.

References

- Chase, W. G. and Simon, H. A. (1973). Perception in Chess. *Cognitive Psychology*, 4, 55–81.
- Christopher, M. (2006). *The Illustrated History of Magic*. New York: Carroll and Graf Publishers.
- Findlay, J. M. (2005). Covert attention and saccadic eye movements. In L. Itti, G. Rees and J. Tsotsos (Eds.), *Neurobiology of Attention* (pp. 114–117). London: Academic Press.
- Friesen, C. K. and Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin and Review*, 5, 490–495.
- Gregory, R. L. (1968). Perceptual Illusions and Brain Models. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 171, 179–296.
- Jackson, R. C., Warren, S. and Abernethy, B. (2006). Anticipation skill and susceptibility to deceptive movement. *Acta Psychologica (Amst)*, 123, 355–371.
- Kingstone, A., Smilek, D. and Eastwood, J. D. (2008). Cognitive Ethology: A new approach for studying human cognition. *British Journal of Psychology*, 99, 317–340.
- Kinlsey, D. (1993). Georges Méliès: Magier der Filmkunst. *Stroemfeld: Roter Stern*.
- Kuhn, G., Amlani, A. A. and Rensink, R. A. (2008). Towards a science of magic. *Trends in Cognitive Sciences*, 12, 349–354.
- Kuhn, G. and Findlay, J. M. (2009). Misdirection, Attention and Awareness. Inattentional blindness reveals temporal relationship between eye movements and visual awareness. *Quarterly Journal of Experimental Psychology*, 1–11.
- Kuhn, G. and Land, M. F. (2006). There's more to magic than meets the eye. *Current Biology*, 16, R950–R951.
- Kuhn, G. and Tatler, B. W. (2005). Magic and fixation: Now you don't see it, now you do. *Perception*, 34, 1155–1161.

- Kuhn, G., Tatler, B. W. and Cole, G. G. (2009). You look where I look! Effect of gaze cues on overt and covert attention in misdirection. *Visual Cognition*, 19, 925–944.
- Kuhn, G., Tatler, B. W., Findlay, J. M. and Cole, G. G. (2008). Misdirection in magic: Implications for the relationship between eye gaze and attention. *Visual Cognition*, 16, 391–405.
- Lamont, P. and Wiseman, R. (1999). *Magic in theory*. Hartfield: Hermetic Press.
- Land, M. F. and Lee, D. N. (1994). Where we look when we steer. *Nature*, 369, 742–744.
- Land, M. F. and McLeod, P. (2000). From eye movements to actions: how batsmen hit the ball. *Nature Neuroscience*, 3, 1340–1345.
- Mack, A. and Rock, I. (1998). *Inattention blindness*. Cambridge, MA: MIT Press.
- Macknik, S. L., King, M., Randi, J., Robbins, A., Teller, Thompson, J. and Martinez-Conde, S. (2008). Attention and awareness in stage magic: Turning tricks into research. *Nature Reviews Neuroscience*, 9, 871–879.
- Parris, B. A., Kuhn, G., Mizon, G. A., Benattayallah, A. and Hodgson, T. L. (2009). Imaging the impossible: An fMRI study of impossible causal relationships in magic tricks. *Neuroimage*, 45, 1033–1039.
- Pinna, B. (2008). *Art and Perception. Towards a Visual Science of Art, Part 1*. Leiden: Brill.
- Rensink, R. A., O'Regan, J. K. and Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 8, 368–373.
- Sharpe, S. H. (1988). *Conjurers Psychological Secrets*. Calgary AB: Hades Publications.
- Triplett, N. (1900). The psychology of conjuring deceptions. *The American Journal of Psychology*, 6, 439–510.
- Tse, P. U. (2004). Mapping visual attention with change blindness: New directions for a new method. *Cognitive Science*, 28, 241–258.

Dreaming as a model system for consciousness research

Antti Revonsuo and Katja Valli

Based on biological realism, that is, the philosophical assumption that consciousness is a real natural biological phenomenon in the brain, we argue that dreaming is a pure form of phenomenality, and suggest that the dreaming brain could be used as a model system for consciousness. The dreaming brain offers the most challenging model system that represents all the theoretically and philosophically interesting features of consciousness. Although difficult to study experimentally, the dreaming brain is the model system that fully preserves all the essential features of consciousness, and cannot be ignored by any theory that aims to explain consciousness.

Sleep and dreaming are the most often naturally occurring altered states of consciousness, involving dramatic changes in overt behavior, physiological responses and subjective experiences. Dreams are experienced by all or nearly all people, and universally across all cultures. Thus, the tendency of the brain to regularly produce subjective hallucinatory experiences during sleep seems to be a part of our inherited biological machinery.

Dreaming as pure phenomenality

Revonsuo (2006) suggests that the dreaming brain constitutes a unique model system that may reveal to us what phenomenal consciousness is and how it is related to external stimuli, observable behaviour, and the underlying brain activity. During dreaming the brain creates a simulated sensory-perceptual world that is independent of concurrent external sources of stimulation. The world-simulation of our dreams takes a similar overall form as phenomenal consciousness does during wakefulness: we experience being one (bodily) self in the center of one unified, animated perceptual world. The simulated world of dreams also shows a rich variability of different experiential contents of consciousness (e.g. sensations, percepts,

emotions, cognitive activities, and social interactions), comparable to waking consciousness (e. g., Strauch and Meier, 1996; for review, see Revonsuo, 2006).

Revonsuo (2006) further argues that dreaming is a pure form of phenomenality, because unlike normal waking consciousness, dreaming occurs in isolation from external stimuli, in the absence of sensory input and motor output. Even though the brain processes both externally and internally generated sensory stimuli during sleep, external stimuli very rarely reach consciousness in the sense that they would be represented in or incorporated into the dream experience. There is a 'sensory input-blockade', a barrier of neural inhibition which guarantees that most sensory stimuli are kept outside of consciousness during dreaming. Similarly, even though motor commands are produced in the brain during dreaming and sent off to reach the appropriate muscles, specific brain stem neurons produce almost complete muscular atonia during REM sleep, blocking the motor commands from reaching their destination. This mechanism is called the 'motor output-blockade'. Therefore, being isolated from the external world as to both the input and the output mechanisms of sensorimotor interaction with the world, dreaming can be regarded as a special and pure form of consciousness, or an 'inner presence' (Revonsuo, 2001; 2006).

The idea of dreaming as pure phenomenal consciousness is rooted in the philosophical assumption that consciousness is a real natural biological phenomenon in the brain. This approach is called 'biological realism' (Revonsuo, 2001; 2006). Biological realism treats consciousness as a biological phenomenon that resides in the brain. Consequently, it follows that subjective consciousness can be given a fully naturalistic explanation, similar to the explanation of other biological phenomena (Revonsuo, 2006; 2010).

Dreaming as a model for investigating consciousness

Revonsuo (2001; 2006) suggests that the dreaming brain could be used as a model system for consciousness, as it has similar overall form as phenomenal consciousness during wakefulness, but is largely isolated from external stimuli, and therefore represents consciousness at its barest. Hobson (1999), in contrast, has proposed that dreaming is a form of delirium, more comparable to a psychotic state than to normal state of consciousness. According to Hobson, dreaming is distinct from waking consciousness due to the chemical imbalance, that is, the modulatory effect between aminergic and cholinergic neurotransmitter systems is inherently different in wakefulness and sleep, especially REM sleep. Thus, Hobson's view seems to be that the dreaming brain can be used as a model system, but not for normal waking consciousness, only for distorted, psychotic consciousness.

In what sense might dreaming serve as a model system in consciousness research? Typically, model systems in science are supposed to be simpler than the actual phenomena they are used to explain, i. e., to manifest the explanandum in a clear form. A model system should allow easy observation and manipulation of the phenomenon. It should be able to reveal quickly the necessary theoretical principles to explain the phenomenon; otherwise it may not be useful. For example, geneticists use the common fruit fly (*Drosophila melanogaster*) as a model organism to investigate the effects of genetic modifications on phenotype, because fruit fly is easy to take care of in laboratories, it has a rapid generation span, it produces numerous offspring, and its genotype is relatively simple and well-known.

Some consciousness researchers use visual awareness as a model system for consciousness, because the anatomical and functional organization of the visual cortex is quite well known and because it can be easily manipulated by external stimuli. This enables the study of for example the neural correlates of changes in experience occurring in the absence of change in external stimuli, such as visual illusions, binocular rivalry, or correlates of unconscious processes such as blindsight. Obviously, visual awareness is a useful model system for empirical and experimental research because it can be easily controlled and manipulated in the laboratory with many kinds of stimuli and tasks. The problem with visual awareness as a model system is that it is difficult to separate stimulus-related (non-conscious or preconscious) processing from the conscious stages of processing the stimulus. And it is equally difficult to separate the responses related to post-perceptual processing (decision making, motor planning) from the stages of pure phenomenal experience of the visual stimulus. Furthermore, one criticism often voiced against visual awareness as a model system is that the subjects are all the time conscious in the experiments, only some restricted aspects or features of the contents of their visual consciousness change. Thus, the results may not reveal the nature of consciousness as such, only the nature of some visual content.

The dreaming brain as a model system has its own strengths and weaknesses which are quite different from visual awareness. Of course, the contents of consciousness cannot easily be manipulated or controlled during dreaming. On the other hand, dreaming as a state of consciousness is both fully stimulus-independent and fully response-independent, and therefore there is no danger to confuse consciousness as such with either physical stimulus effects or with premotor planning or verbal responses. Dreaming is “pure” consciousness in this sense. Furthermore, in the sleeping brain, there is variation both in the overall state of consciousness (dreaming vs. dreamless sleep) as well as in the particular contents of dreaming (visual, auditory, emotions, bodily self, etc.) and in the overall complexity of the experience (simple static mentation vs. simple dreaming vs. complex dreaming). Thus, these patterns of variation at the phenomenal level should

be accompanied by corresponding patterns at the level of the directly underlying neural mechanisms of consciousness.

Moreover, dreaming is a maximally subjective phenomenon whose presence or precise contents cannot be derived by external observers from any objective measurements concerning the behavior, physiology, or brain activity of the sleeping person. The dreaming brain thus offers the ultimate challenge for the science of consciousness, a test case as to whether the science of consciousness is able to solve the deepest philosophical and theoretical problems of consciousness: the Explanatory Gap, the Hard Problem, the Other Minds problem, and the problem of objectively measuring or detecting consciousness. Only if one fine day it is possible to explain how brain activity exactly produces the phenomenal contents of consciousness has the Hard Problem and the Explanatory Gap been solved. And only if one fine day it will be possible to tell with certainty, without ever asking the subject, whether the subject who is sleeping is also dreaming and what the content of his dream is like, has the problem of objectively measuring phenomenal consciousness been solved.

Thus, whereas visual awareness offers perhaps the easiest model system for consciousness research, a system that is easily manipulated and controlled, but that represents only some restricted features of consciousness, the dreaming brain offers the most challenging model system that represents all the theoretically and philosophically interesting features of consciousness. While studies of visual awareness offer some relatively easy ways to experiment with consciousness, they may at the same time oversimplify the phenomenology of consciousness to some very limited aspect of perception. Studies of the dreaming brain by contrast involve studying consciousness as an overall world-simulation system in the brain where the whole richness of phenomenal experience is manifested at the same time as no external stimuli or motor responses interfere. Studying the dreaming brain may be more difficult experimentally, but at least we are then studying the real thing as it is: consciousness as a holistic, dynamic system in the brain rather than reducing it to some simple visual sensation or percept only.

Problems and accomplishments

We already have systematic data about the statistical features of dream content. The problem is to correlate dream content (and the presence vs. absence of dreaming) with brain activity. The problem is that we do not know when exactly dream experiences occur. In one extreme, there is the amnesia hypothesis according to which we dream all through the night, but sometimes we simply fail to remember our dreams. In the other extreme is the false-memory hypothesis, according to

which we construct dream reports only when we wake up from sleep and seem to remember having had a dream; in reality our sleep was dreamless and we only have false memories, confabulated at awakening. It is difficult to conclusively refute either of these directly opposite hypotheses, which is only one more indication of the fact that we have no objective manner to confirm either the presence or the absence of dreaming.

If we take the data produced in dream research seriously, however, dreaming could be helpful in solving the binding problem. The binding problem refers to how the unity of conscious perception is brought about by the distributed activities of the central nervous system (Revonsuo and Newman, 1999). Bizarreness in dreams can be reconceptualised as a failure in binding elements coherently together, and dream bizarreness could be used to explain under what conditions binding fails during waking consciousness, and to detect which underlying neural mechanisms participate in binding. This notion has led to one interesting study about bizarreness and mismatches in binding the internal and contextual features of dream characters together (Revonsuo and Tarkko, 2002).

Another theoretically relevant contribution from dream science to consciousness studies comes from an evolutionary psychological theory of the function of dreaming, the threat simulation theory (Revonsuo, 2000). The theory originated from questions such as: why is dream consciousness so negatively toned, and why do traumatic events have such a powerful impact on subsequent dreaming? In the end, a new theory was formulated based on the phenomenological features of dreams. The theoretical contribution of the threat simulation theory does not lie only in providing an evolutionary biological explanation for the function of dreaming, but is philosophically significant because it attaches causal powers to phenomenal dream consciousness, and thus gives a significant functional role for the dreaming brain.

Phenomenal consciousness does seem to appear in its purest form during dreaming, that is, it is created by the brain in the absence of input and output. It is in itself ecologically valid, and needs no experimental manipulation to be created. Yet, for consciousness research, dreaming may not be the clearest and simplest model system, comparable to visual awareness. But the dreaming brain is the model system that fully preserves all the essential features of consciousness: its holistic unity as a dynamic world-simulation, its richness and variability of content, its subjectivity as independent of stimuli, motor output, or detection by objective research instruments. It is the model system that presents the ultimate challenge for consciousness research: If we can explain how phenomenal dream experience arises from the sleeping brain, we will thereby have provided an explanation to all the most difficult problems that consciousness in general poses to science.

Conclusion

Regardless, any theory that aims to explain consciousness (with whatever model system) must also provide an explanation for the nocturnal form of consciousness. This kind of universal, cross-generationally occurring, most naturally induced altered state of consciousness cannot be ignored, no matter how difficult it is to explain. There is no doubt that consciousness, whether we talk about normal waking consciousness or hallucinations during wakefulness or dreaming, is internally generated and resides in the brain. Therefore, theories of consciousness that contradict the basic phenomenological or neurophysiological findings of dream research, or do not account for dreaming, are most likely incomplete or completely inaccurate.

References

- Hobson, J. A. (1999). *Dreaming as delirium. How the brain goes out of its mind*. Cambridge, MA: MIT Press.
- Revonsuo, A. (2000). The reinterpretation of dreams: An evolutionary hypothesis of the function of dreaming. *Behavioral and Brain Sciences*, 23, 877–901.
- Revonsuo, A. (2001). On the nature of explanation in neurosciences. In P. Machamer, R. Grush and P. McLaughlin (Eds.), *Theory and method in the neurosciences* (pp. 45–69). Pittsburgh: University of Pittsburgh Press.
- Revonsuo, A. (2006). *Inner presence: Consciousness as a biological phenomenon*. Cambridge, MA: MIT Press.
- Revonsuo, A. (2010). *Consciousness, the science of subjectivity*. London: Psychology Press.
- Revonsuo, A. and Newman, J. (1999). Binding and Consciousness. *Consciousness and Cognition*, 8, 123–127.
- Revonsuo, A. and Tarkko, K. (2002). Binding in dreams: The bizarreness of dream images and the unity of consciousness. *Journal of Consciousness Studies*, 9, 3–24.

Lucid dreaming and the bimodality of consciousness

Allan Hobson and Ursula Voss

In this essay, we develop our working hypothesis that consciousness in primates and humans is a state-dependent commodity that has at least two expressions. Waking and dreaming are two such states that differ in conscious awareness. In both states, we are perceptive and emotional. Whereas in dreaming, our thoughts are delusional, however, waking consciousness potentiates volition and reflection. We propose that dreaming is a state of primary consciousness, while usually, secondary consciousness is reserved to waking. Lucid dreaming is an extraordinary state with elements of both waking and dreaming and both primary and secondary consciousness. It is a rare but very real condition which is a promising tool in the study of the brain basis of consciousness.

Reflective consciousness is the phenomenon that catapults mankind to the forefront of evolution. Whereas most mammals possess a primitive (or primary) form of consciousness, humans and primates also possess a higher-order (or secondary) form of consciousness which provides the prerequisite for abstract thought and volition.

Consciousness must therefore be at least bimodal. Research on sleep and dreaming has documented the state-dependency of this bimodality which is expressed in the differential physiology and phenomenology of the dreaming and waking brain and our subjective awareness of these states. Whereas secondary consciousness is usually accessed only in waking, dreaming may be understood as primary consciousness. In lucid dreaming the sleeper gains cognitive insight into the fact that he is dreaming instead of delusionally believing himself to be awake. Lucid dreaming, therefore, presents us with the unique opportunity to monitor how secondary consciousness arises from a state that is usually restricted to primary conscious awareness.

The bimodality of consciousness

Until recently, waking was regarded as the state of consciousness, par excellence. Dreaming was relegated to a subservient and disruptive role as the repressed unconscious of Sigmund Freud and his followers. But as previously argued (Hobson, 2009a), dreaming is mostly forgotten (and hence amnesic) but not really unconscious. When we dream, we are, if anything, hyperperceptive and hyperemotional. It is this augmentation of conscious state features that led Leonardo da Vinci to wonder, “Why does the eye see a thing more clearly in dreams than it does when awake?” At the same time, other, secondary, conscious state features of waking consciousness, including memory, are in abeyance. We *can* spontaneously remember some dreams. And we can remember many more if awakened by autosuggestion, by an accomplice, or by a technician in a sleep lab. Dreaming may therefore be regarded as a second state of consciousness, quite different from, but quite complimentary to, waking consciousness.

These facts and considerations indicate that consciousness is, at least, bimodal. It may, in fact, be multimodal but for the purposes of this essay, we choose to focus on its indubitable bimodality.

The bimodality of consciousness converges closely with Gerald Edelman’s (2003, 2005) useful distinction of *primary and secondary consciousness*. While Edelman assumes that *primary consciousness* may be experienced by most birds and mammals, he ascribes *secondary consciousness* to primates, especially humans. Primary consciousness is a state that allows for simple awareness, including perception and emotion. Humans and primates have developed a superior form of consciousness referred to as secondary consciousness. This involves awareness of awareness which requires volition, rational thought, and self-reflection.

For strategic scientific reasons, we suggest that dreaming is like primary consciousness. In dreaming, our consciousness is strongly perceptive and highly animated yet it lacks the characteristics of secondary consciousness.

The distinction between the two modes of conscious experience begs comparison with Freud’s primary process (which he ascribed, as we do, to dreaming) and secondary process (which he ascribed, as we do, to waking). But whereas Freud conceived of the two processes as fundamentally antagonistic, we view them as fundamentally cooperative.

Evidence for bimodality: Consciousness in waking and in dreaming

What is the evidence for bimodality? What is its mechanism? And what is its function?

- In waking, my perceptions are strongly shaped by external inputs; in dreaming equally or even more vivid perceptions are internally generated.
- In waking, I experience my will as free and make elaborately planned motor acts in keeping with my perceptions; in dreaming, I am always motorically active and have the strong illusion of movement through a real space, although my body is largely immobile as I lie asleep in my bed.
- In waking the unities of time, place and person are preserved; in dreaming, I move from place to place, from era to era, and interact with one person after another while my identification of those places, eras and persons may be vague, inaccurate, and arbitrary. In dreaming it is as if the unities of time, place and person had been dissolved.
- In waking, I can not only orient faithfully but I can think, plan and review rationally; in dreaming my thought processes are greatly impaired. Not only do I not know what is going on when I dream but I don't know that I don't know what is going on.
- In waking, I am capable of moderating my emotions; in dreaming I often experience uncontrolled passion; and finally
- In waking I remember some things very well; in dreaming I forget almost everything.

These phenomenological differences are greatest when one compares alert waking to REM sleep. Intermediate and mixed states exist but will be ignored in this essay.

Mechanism: The AIM model

It is now quite clear that dreaming is caused by brain activation in sleep and that activation, like the activation of waking, is mediated by the pontomesencephalic brain stem.

The Activation function, *A*, of the AIM Model (Hobson et al., 2001; Hobson, 2009a) makes no distinction between waking and dreaming although quantitative EEG studies to be discussed below, do show differences.

According to AIM, the bimodality of waking and dreaming is caused by two other processes, also of brainstem origin.

One is the second AIM factor, *I*, which distinguishes between Input-output gates: these gates are open in waking, but are closed in sleep.

The third AIM function, factor *M*, posits that the Mode of information processing is altered such that waking is strongly and dreaming only very weakly, aminergic. The two conscious states of waking and dreaming are thus chemically distinct.

In summary, the brain is activated, online, and aminergically modulated in waking; in dreaming the brain is also activated but it is offline and aminergically demodulated. The precise cellular and molecular mechanisms of the bimodality are beginning to be understood but are beyond the scope of this essay (Hobson et al., 2001).

Function

What is the relationship between the conscious states of dreaming and waking and what does that relationship tell us about consciousness? One answer to this question is evolutionary: animals first developed a relatively simple system for mediating their perceptions, emotions, and memories. This was, and still is, the primary consciousness system which is all that birds and sub-primate mammals have to work with. The primary consciousness system is activated in REM sleep in all animals that possess its advantages. As the brain continued to evolve, and greatly elaborated its thalamocortical circuitry, it became possible for animals to add secondary aspects of consciousness to their waking repertoire.

The enormous advantages of these capabilities are obvious. But there is a problem: how are primary and secondary mechanisms kept separate and how do they conspire to interact positively?

The state dependency of consciousness

Separation is conferred by state differentiation. This is seen clearly in the differential phenomenology and mechanism of waking and dreaming. But note, both subprimate and primate animals enjoy the benefits (and risks) of a primary consciousness system. There is strong evidence that primary consciousness co-develops with REM sleep and constitutes a virtual reality generator for the brain (see, again, Table I). When the REM generator is activated in sleep, dreaming occurs. An implication of this assertion is that the brain contains within its own intrinsic operations, representations of self, of space, of movement and of feeling. In waking it is important to be able to damp this system so that cognition can be kept in register with external world parameters. All mammalian animals can do this but the evolution of secondary consciousness has rendered this task more difficult and its failure more risky. Mental illness may well be a manifestation of this risk.

An alternative view would claim that state differentiation is not necessary and that we are able to maintain different modes of consciousness in a single state. For example, the default mode model (Raichle et al., 2001) suggests that consciousness differs, depending on whether we are attending to either inwardly

(default mode) or outwardly directed thoughts (information processing mode), both of which alternate constantly in waking. This idea is in perfect keeping with factor I of the AIM model, and we are aware that several sub-states exist that will also require identification and lead us to a multimodal model of consciousness.

However interesting, these considerations are not helpful in explaining phylogenetic differences in conscious awareness. For the same reason, they are not compatible with the primary/secondary consciousness distinction that we emphasize in our bimodality thesis.

It is important to recognize that consciousness must be regarded as a *state*, and not as a specific brain operation. The respective mode of consciousness provides the background against which certain functions may be executed. Consciousness must be regarded as the brain's *capacity* to perform inwardly or outwardly directed cognitive functions.

When the primary REM system escapes from control, waking consciousness becomes impossible. Cognition then becomes dominated by internal perceptions (visual hallucinations), disorientation, recent memory loss, and confabulation. This is the organic mental syndrome that is often seen in patients suffering from delirium (Hobson, 2000). From a theoretical point of view, even subprimate animals should run this risk but it is possible that important factors, such as drugs and alcohol uniquely predispose humans to this problem. In any case, subprimate animals *do* hold the REM-primary consciousness generator in check as they navigate their relatively simpler waking worlds.

A second biological factor makes this account plausible and attractive. During the early development of those mammalian species that are conscious, REM sleep occurs. It may even occur in utero but is certainly prominent directly after birth. In kittens, it plummets sharply as eyes open waking comes to take its place (Herman et al., 1991).

Besides providing the brain with a virtual world model, REM sleep guarantees a crucial adaptive capability, homeothermy. In order to have a complex brain function normally it is critical to control its temperature. REM confers that capability (Rechtschaffen et al., 1989).

Lucid dreaming, an experiment of nature

If there really are two brain systems which conspire to engineer waking and dreaming consciousness then it should be possible to tease them apart and identify the circuitry that makes the two states of consciousness so different. Such a physiological dissection might also help us understand the mutual collaboration of dreaming and waking and to model its occasional degeneration.

When most subjects dream, they assume, wrongly, that they are awake. This delusional mistake has at least two sources: one is that waking and dreaming consciousness are phenomenologically similar as well as stunningly different from each other. The differences are unnoticed because dream perception, fictive movement, and emotion are so strong as to be pre-emptive and because cognitive discrimination suffers from the diminution in memory, orientation, reasoning power and amnesia of dreaming.

We sometimes suspect that our dream consciousness is not waking but usually brush such criticism aside. When, through training or spontaneous change we become convinced that we are not awake but, rather, dreaming we enter the remarkable state called lucid dreaming. Lucid dreaming is difficult to attain and to maintain, probably because state differentiation is adaptive and therefore the brain is engineered to be in one state or the other but not both simultaneously. The occurrence of lucid dreaming thus presents consciousness research with a unique opportunity to understand the mechanism of primary and secondary consciousness and begin to explore the functional significance of their interaction.

This challenge is met via the development of new tools for cognitive science including quantitative EEG (qEEG), functional brain imaging (fMRI) and electromagnetic brain stimulation (EMS). Preliminary results using the first two of these three methods are promising.

Laboratory studies

Physiologic correlates of lucid dreaming were first objectively studied in the laboratory by K. M. Hearne (1978) and Steven LaBerge (1980, see also Holzinger et al., 2006). However, methodological difficulties and concerns (Tart, 1979; Hobson, 2009b) discredited the scientific status of these studies and moved lucid dreaming into esoteric realms. LaBerge convincingly demonstrated, however, that lucidity consistently arose out of REM sleep and that it was possible for subjects to signal out of REM that they were lucid by making a series of voluntary eye movements. An unresolved question was whether the lucid subjects were still in REM sleep or whether they had transcended into waking. Our own study revealed that the brain in lucid dreaming is both awake and asleep at the same time (Voss et al., 2009). This paradox not only sheds light on the brain mechanisms of bimodal consciousness but shows that the REM sleep vs. waking controversy can be resolved by the assumption of a both-and rather than an either-or position.

In fact, the brain may never be in one and only one state; instead dissociations abound as Mark Mahowald has so cogently argued (2009).

When we studied lucid dreaming, the greatest experimental obstacle was that lucidity is a very fragile state. It is, therefore, difficult to observe in a laboratory setting. Our subjects were sensitive to noise and less able to perform plot control when sleeping in the laboratory compared to sleeping at home. Although lucid dreaming can be trained, it requires effort and determination. Accordingly, in our first EEG study, we only recorded 3 episodes that were clearly objectifiable as spontaneous lucid dreams. Despite the low number of observations, we found that when subjects became lucid, they shifted their EEG power (especially in the 40 Hz range) in frontal regions of the brain (see Figure 1). These changes were statistically significant.

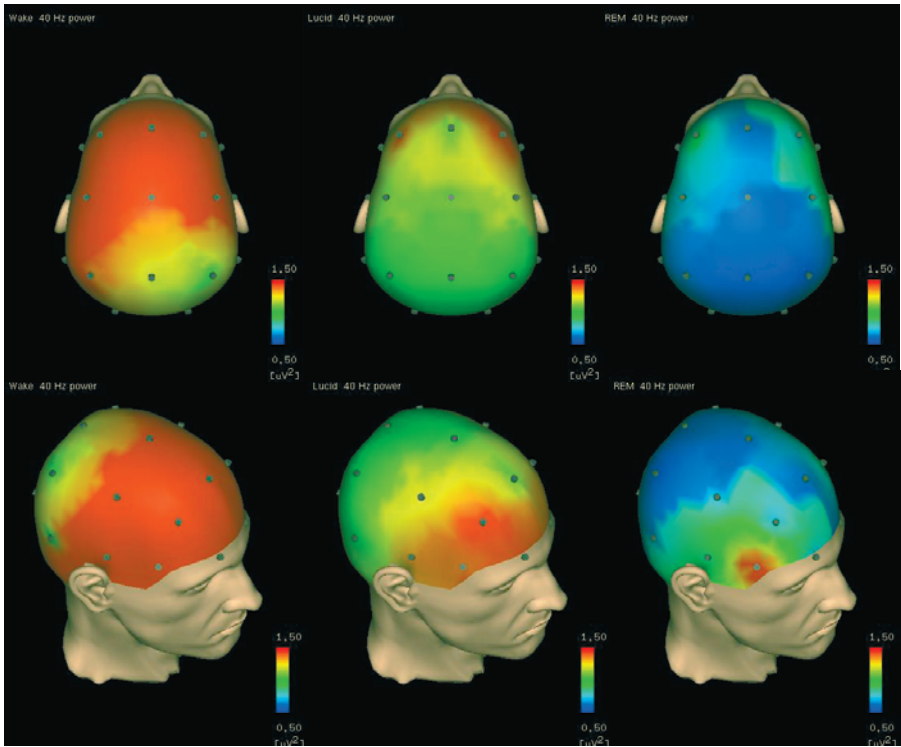


Figure 1. Standardized 40 Hz power (36–45 Hz) in waking (left), lucid dreaming (middle) and REM sleep (right frame) in a single subject. Topographic images are based on movement-free EEG episodes and are corrected for ocular artifacts. For each state, power values are averaged across the respective episode. Compared to REM sleep, lucid dreaming shows increased 40 Hz activity across the entire cortex. The increase is strongest in frontal regions. Compared to waking, frontal and parietal regions are similarly activated in lucid dreaming.

We also found a marked increase in cortical networking, as expressed by enhanced coherences, especially between occipito-parietal and frontal areas of the brain (see Figure 2).

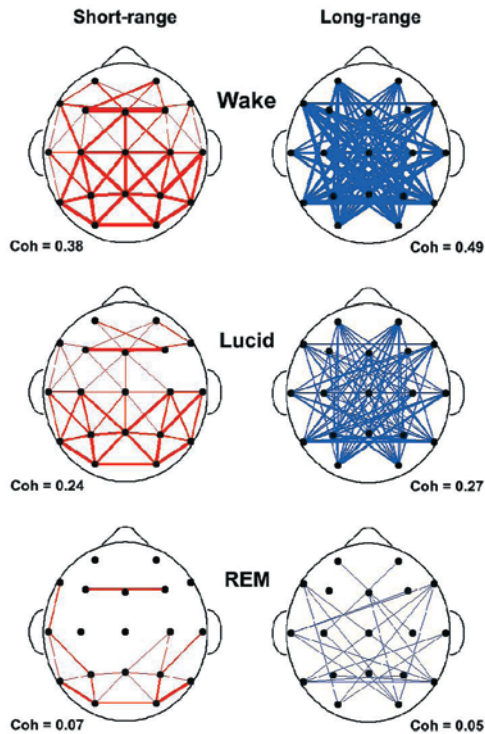


Figure 2. Short and long range coherences obtained for waking (top), lucid dreaming (middle) and REM sleep (bottom). Coherences are averaged across electrode pairs in 4 s epochs. Short-range (55 pairs) was defined as less than 10 cm and long-range (65 pairs) as larger than 15 cm inter-electrode distance. Coherences are lowest in REM sleep and strongly enhanced in lucid dreaming.

We emphasize that the changes in cortical activation we have described are a consequence of pre-sleep auto-suggestion indicating that in humans even primary consciousness, which is largely automatic, is partially subject to volitional force. This fact has important therapeutic implications as well as telling us that waking (secondary) consciousness may differ from (primary) dreaming consciousness owing to increases in the power and coherence of the frontal lobes of the brain.

Recent functional MRI results reported to us by our Munich colleagues support this conclusion by revealing that dream lucidity is correlated with increased

activation of the cortical areas that are thought to mediate the features of secondary consciousness. The selectively activated brain regions include the precuneus, the frontopolar cortex, the dorsolateral prefrontal cortex, the temporal cortex and the inferior parietal lobules (Dresler et al., p.c.)

Comparison of recent results to previous imaging studies

Imaging studies of normal sleep have consistently shown that compared to waking, REM sleep is characterized by diminished activation of the dorsolateral prefrontal cortex (DLPFC) (Maquet et al., 1996; Muzur et al., 2002; Nofzinger et al., 1997). Since the DLPFC is thought to be the site of executive ego function, it has been suggested that the loss of volition, self-reflective awareness, and insight that is typical of normal dreaming, may be related to DLPFC inactivation. In lucid dreaming, these psychological functions are regained leading to the prediction that lucid dreaming involves reactivation of the DLPFC.

Lucid dreaming as a bimodal state of consciousness

We interpret our recent results and previous findings as follows. There are two cardinal states of consciousness, waking and dreaming. Each has distinctive phenomenologic and physiological features. Their co-existence in all animals that evince REM sleep indicates shared brain mechanisms and functions for these states. We have elsewhere argued that the two states are mutually reinforcing (Hobson, 2009a).

In the current paper we describe our experimental results in terms of our theory of bimodality. We have shown that lucid dreaming is an unusual but informative state with features of both waking and dreaming. As such its physiological substrate demonstrates the simultaneous co-activation of brain circuits mediating REM sleep and waking.

These observations and our interpretation of them have far-reaching ramifications for theories of consciousness. For example, the widely noted and much appreciated unity of consciousness is called into question. Not only is consciousness bimodal but at least two of its modes can be simultaneously activated, experienced and observed. As such, lucid dreaming is a hybrid state in which one part of the brain (and its subjective self) observes and controls another part of the brain (with its own subjective self). It is in this sense that we allude to the famous split brain studies of Sperry (1961) and Gazzaniga (1989); we suggest that brain-mind

splitting is not only anatomical and post-surgical (as in their cases) but also physiological and post-training (in our lucid dreamers).

Multiple personality disorder, hypnosis, and major mental illnesses like schizophrenia are all better understood via the recognition that even normal sleep is characterized by surprising bimodality. The possibility that the observed bimodality reflects profound and adaptive functionality serves to extend and naturalize our understanding of otherwise puzzling conditions of the brain-mind.

Summary and conclusions

Consciousness is not a unitary phenomenon and it is not universal. Whereas most mammals possess a primary level of conscious awareness, secondary consciousness (enabling rational thought and volition) arises in humans, possibly primates and birds. Consciousness is at least bimodal in humans and primates which are known to possess secondary consciousness. The bimodality is usually maintained through state-differentiation: REM sleep dreaming is dominated by primary consciousness and waking by secondary consciousness which tends to make the underlying primary features less easily appreciated. In dreaming, the brain is highly activated, and aminergically demodulated, while the external world is excluded. Thought is irrational and volition strongly attenuated. In waking, primary consciousness is complimented by secondary consciousness, giving way to rational thought and self-reflection.

In contrast to the competition suggested by the Freudian school, primary and secondary consciousness are not necessarily antagonistic or dangerous to each other. Rather, primary consciousness is seen as the necessary cognitive base on which secondary consciousness is built. Dreaming, thus, represents the cornerstone of higher-level consciousness. Experimental proof of this dependency comes from studies of lucid dreaming, in which we can monitor the emergence of rational thought and volition in a brain state usually dominated by primary consciousness.

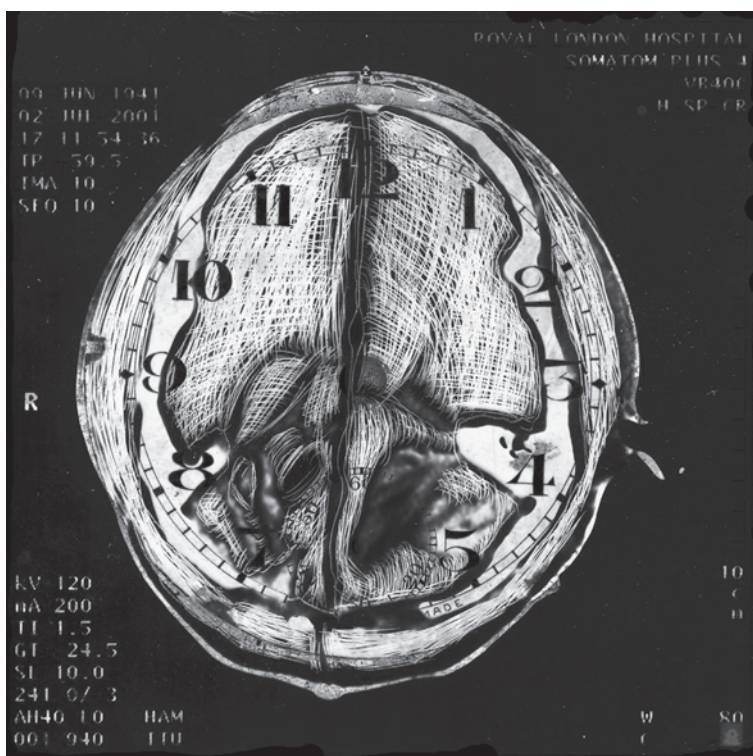
References

- Dresler, M., Koch, S., Wehrle, R. et al. Neural correlates of consciousness visualized by sleep imaging. Personal communication, November 10, 2008.
- Edelman, G. M. (2005). *Wider than the sky: a revolutionary view of consciousness*. London: Penguin Press Science.
- Gazzaniga, M. S. (1989). Organization of the human brain. *Science*, 245, 947–952.

- Hobson, J. A. (2000). *Dreaming as Delirium: How the Brain Goes Out of Its Mind*. Boston: MIT Press.
- Hobson, J. A. (2009a). REM sleep and dreaming: Towards a theory of protoconsciousness. *Nature Reviews Neuroscience*, 10, 803–813.
- Hobson, J. A. (2009b). The Neurobiology of Consciousness: Lucid Dreaming Wakes Up. *International Journal of Dream Research*, 2, 41–44.
- Hobson, J. A., Pace-Schott, E. F. and Stickgold, R. (2001). Dreaming and the brain: Toward a cognitive neuroscience of conscious states. *Behavioral and Brain Sciences*, 23, 793–842.
- Hearne, K. (1978). *Lucid dreams: An electro-physiological and psychological study*. University of Liverpool, England, Dept. of Psychology: PhD thesis.
- Herman, M. D., Denlinger, S. L., Patarca, R., Katz, L. and Hobson, J. A. (1991). Developmental phases of sleep and motor behaviour in a cat mother–infant system: A time-lapse video approach. *Canadian Journal of Psychology*, 45, 101–114.
- Holzinger, B., LaBerge, S. and Levitan, L. (2006). Psychophysiological correlates of lucid dreaming. *Dreaming*, 6, 88–95.
- LaBerge, S. (1980). *Lucid dreaming: An exploratory study of consciousness during sleep*. Stanford University: Doctoral dissertation.
- LaBerge, S., Nagel, L. E., Dement, W. C. and Zarcone, V. P. (1981). Lucid dreaming verified by volitional communication during REM sleep. *Perceptual and Motor Skills*, 52, 727–32.
- Mahowald, M. W. (2009). What state dissociation can teach us about consciousness and the function of sleep. *Sleep Medicine*, 10, 159–160.
- Maquet, P., Peters, J. M., Aerts, J., Delfiore, G., Degueldre, C., Luxen, A. and Franck, G. (1996). Functional neuroanatomy of human rapid-eye-movement sleep and dreaming. *Nature*, 383, 163–166.
- Muzur, A., Pace-Schott, E. F. and Hobson, J. A. (2002). The prefrontal cortex in sleep. *Trends in Cognitive Sciences*, 6, 475–481.
- Nofzinger, E. A., Mintun, M. A., Wiseman, M., Kupfer, D. J. and Moore, R. Y. (1997). Forebrain activation in REM sleep: An FDG PET study. *Brain Research*, 770, 192–201.
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A. and Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 676–682.
- Rechtschaffen, A., Bergmann, B. M., Everson, C. A., Kushida, C. A. and Gilliland, M. A. (1989). Sleep deprivation in the rat: X. Integration and discussion of the findings. *Sleep*, 12, 68–87.
- Sperry, R. W. (1961). Cerebral organization and behavior – Split brain behaves in many respects like 2 separate brains, providing new research possibilities. *Science*, 133, 1749–1757.
- Tart, C. T. (1979). From spontaneous event to lucidity: A review of attempts to consciously control nocturnal dreaming. In B. B. Wolman (Ed.), *Handbook of dreams* (pp. 226–268). New York: Van Nostrand.
- Voss, U., Holzmann, R., Tuin, I. and Hobson, J. A. (2009). Lucid dreaming: A state of consciousness with features of both waking and non-lucid dreaming. *Sleep*, 32, 1191–1200.

SECTION III

Psychopathologies and therapies



Brain 12, Cogito Ergo Sum (detail). Susan Aldworth (2006) giclee print, 35×35 cm

Why depression feels bad

Mark Solms and Jaak Panksepp

We believe that conscious mental phenomena (such as feelings) are not epiphenomenal to the workings of the brain. Feelings evolved for good biological reasons; they make specific, concrete contributions to brain functioning. Notwithstanding all the philosophical complexities, therefore, the non-conscious/conscious interactions that are the focus of this book are, in our view, causal interactions. To marginalize consciousness in relation to what is ultimately a dualistic scientific understanding of how the brain works is likely to lead us astray. We illustrate this view by trying to address the question: why does depression feel bad?

When Sigmund Freud first argued that the mind is not synonymous with consciousness, he was roundly criticised, mainly on philosophical grounds. Subsequent empirical findings have, however, strongly supported his view that many if not most mental functions do not require consciousness to operate efficiently and effectively. In fact, the evidence for this view is now so overwhelming that the converse question is frequently asked: why do we need consciousness at all? This question deeply haunts contemporary efforts (like this timely book) to explore and explain the many complexities that surround non-conscious/conscious interactions in the brain.

It is therefore interesting to note that in a posthumously-published outline of his life's work Freud (1940, p. 157) acknowledged that consciousness was *the most unique characteristic* of the part of nature that we call the mind – 'a fact without parallel'. The fact of consciousness, Freud wrote, 'defies all explanation and description'. He continued: 'Nevertheless, if anyone speaks of consciousness we know immediately and from our most personal experience what is meant by it'. He then added a disparaging remark to the effect that 'one extreme line of thought, exemplified in the American doctrine of behaviourism [which was just then coming to prominence], thinks it possible to construct a psychology which disregards this fundamental fact!'

Behaviourism

It is well known why behaviourists wanted to construct a science of the mind that disregarded its most unique characteristic. Consciousness cannot be observed externally; it is not amenable to direct objective scrutiny. Consciousness is for that reason an embarrassment to science, the ideal of which is objective fact over subjective experience. The behaviourists, who wanted to treat the mind as if it were no different from any other part of nature, therefore ruled consciousness out of court, and limited scientific psychology to the study of the objectively observable *outputs* of the mind – to the study of behaviour. Observable experimental manipulations ('stimuli') could then be used to discover the causal mechanisms of behavioural 'responses'. In this way, the intervening variables (conceptualised as the laws of *learning*) became the only valid objects of psychological science.

Not surprisingly, a school of thought predicated on the assumption that the mind consists in nothing but learning, and disregards all the other mental phenomena that we 'know immediately and from our most personal experience', was doomed to failure. To deny the causal influence on behaviour of conscious states (like feelings) is to deny the obvious. If one says: that person committed suicide because he could not stand the pain any longer; one is describing the simple causal power of that person's feelings. If one were to try to re-phrase such causal statements so as to exclude the feelings, one would be doing violence to the obvious facts.

Thankfully, therefore, in the psychology of the last quarter of the twentieth century, realism triumphed over fundamentalism, and consciousness found its way back into science. Even though consciousness still cannot be observed directly, or objectively, today we are nevertheless willing to acknowledge its existence in our experimental subjects, and on this basis to infer the causal mechanisms by which conscious states influence behaviours – in much the same way as the behaviourists were willing to infer the causal mechanisms of learning... or are we?

The mechanisms of consciousness may be ontologically equivalent to those of learning (or anything else) but the *mechanisms* of consciousness differ in fundamental respects from consciousness *itself*. Mechanisms of all kinds are abstractions, derived from experience; they are not experiences themselves. The mechanisms of consciousness, like all other mechanisms, therefore present no special problems for science; they, too, can be described from an objective standpoint, from the third-person point of view. But this excludes the 'fundamental fact' of consciousness, namely that we *experience* it directly. Is consciousness not perhaps still an embarrassment to science; do we not perhaps still think it possible to construct a psychology which disregards the causal role of this uniquely subjective characteristic of the mind – the fundamental aspect of this part of nature?

Cognitive neuroscience

It is, in our view, no accident that the apparent re-admittance of consciousness to psychology coincided with advances in the neurosciences which made it possible to study the *physiological correlates* of almost any mental state. By shifting the focus of our research efforts to the physical *correlates* of consciousness, we were able to pay lip service to its existence without having to trouble ourselves too much with its intrinsically subjective nature – with the original source of the embarrassment. Small wonder, then, that so many behaviourists made such a seamless transition to the new paradigm.

As Freud (ibid.) put it:

‘there would thus be no alternative left to assuming that there are physical or somatic processes which are concomitant with the psychical ones and which we should necessarily have to recognize as more complete than the psychical sequences [... Then] it of course becomes plausible to lay the stress in psychology on those somatic processes, to see in *them* the true essence of what is psychical’.

To seek the essence of what is mental in something which lacks its most unique psychic property is surely to look in the wrong place. But this does not mean that we must abandon reality. Nor does it mean that the brain is the wrong place to seek an understanding of consciousness. It means only that we must admit that consciousness actually exists, that it is a property of nature, that it is a property of the part of nature called the brain, and that this property is no less real and no less causally efficacious than any other natural properties. This in turn means that we must recognise that the brain is not *quite* the same as every other part of nature. The brain has some special properties, and central among these is consciousness. As a consequence of it being conscious, the brain behaves differently from most other things, even from other bodily organs.

As far as we can tell, despite appearances, these views are still not generally accepted, or at least they are not generally incorporated into the current theoretical paradigms of cognitive neuroscience. In fact the very power of cognitive neuroscience seems to be that it treats the organ of the mind as if it were no different from any other bodily organ or piece of matter, indeed from any other complex mechanism – living or dead.

Biological psychiatry

The baneful consequences of this continued neglect of the ‘fundamental fact’ of consciousness have been more evident in the field of biological psychiatry than

in cognitive neuroscience in general. This is perhaps not surprising, because psychiatry is all about feelings.

In psychiatry today, if one says: the patient committed suicide because he could not stand the pain any longer, one seems to mean: the patient *thought* he was committing suicide because he could not stand the pain any longer, but *really* he was committing suicide because his serotonin levels were depleted (or something like that). The point is: what the patient says, thinks or feels may be left out of our scientific account; the feelings evidently are not really part of the causal chain of events. They are just a layperson's translation of the actual state of affairs in the brain. This, in our view, is doing violence to the facts. In our view, the feelings are a fundamental part of the actual state of affairs in the brain.

We shall now illustrate these principles with reference to a particular problem in modern psychiatry, namely: *why depression feels bad*.

In fact, this problem is not even posed in psychiatry. It is not posed because what depression feels like does not seem to matter in contemporary psychiatric science. This is evidently because feelings in general do not matter. What matters is the physical *correlates* of the feelings. This approach, in our view, is based on a serious misconception of how the brain works, which will almost inevitably lead to equally serious mistakes.

In their haste to avoid the embarrassingly subjective phenomena of depression, psychiatric researchers have in recent decades focused on all sorts of things that correlate with depression, or facilitate it, or contextualise it – and the neural mechanisms of those things – rather than the feelings of depression itself.

The main focus of depression research for the past three decades has been the neurophysiological mechanisms of norepinephrine and serotonin depletion (Schildkraut, 1965; Harro and Oreland, 2001), including the neurotrophic effects of this depletion (Koziak et al., 2008), the neuroendocrinological mechanisms of stress (which has similar neurotrophic consequences; De Kloet et al., 2005), the neuroimmunological equivalents of these mechanisms (McEwen, 2007), their interactions with sleep mechanisms (Zupancic and Guilleminault, 2006), their genetic underpinnings (Levinson, 2006), and so on.

These research programmes have evidently been followed because the mechanisms of serotonin depletion (and its cognates) are eminently tractable scientific problems – notwithstanding the fact that they have little to do with actually researching depressive feelings. The reason these programmes have been followed cannot be because anyone seriously thinks that depressive feelings (let alone major depression) are actually *caused* by low levels of serotonin. There is no direct evidence for that notion. In fact, it is well established that experimental depletion of brain serotonin does not cause depression in normal people (Delgado et al.,

1990). Nor was there ever any sustained line of research to believe that serotonin *would* play any such specific causal role in depressive mood. Serotonin is an all-purpose modulator of moods and emotions, not only of depressive ones (Berger et al., 2009). It is probably for this reason that SSRIs are used to treat not only depression but also a host of other emotional troubles, such as panic attacks and obsessive compulsive disorder. This is also probably the reason why SSRIs do *not* work in so many cases of depression, and why they work only partially or temporarily in the vast majority of cases (cf. STAR*D findings). The same applies to the various cascades associated with serotonin depletion: stress or inflammation or hippocampal shrinkage (Miller et al., 2009). None of these things has a unique and specific causal relationship to depression. They are too general; ‘too much’ of an explanation. Their main attraction is only that they are scientifically tractable and therefore scientifically respectable mechanisms.

In summary, it is clear that although the mechanisms of serotonin depletion and its cognates correlate with or facilitate or contextualise depression, something else – something far more specific – must be the actual causal mechanism of depression. We suggest that this ‘something else’ most likely has something to do with the brain mechanisms that actually generate depressive *feelings*.

Depression itself

Our reason for suggesting this is the fact that the clinical phenomenology of depression is characterized above all else by a complex of feelings: low mood, low self-esteem, loss of motivation and energy, loss of pleasure in the world, and so on. Is this feeling complex not the most obvious place to seek the essential nature of depression? And dare we ask whether this constellation of feelings *means* anything? It is after all in the essential nature of feelings that they mean something. It would be entirely normal and reasonable for all of us (even us scientists) to ask – outside of our scientific work – what it might mean when somebody says that they feel down, bad, defeated, useless; that they have lost all hope for themselves, lost all interest in other people, and so on. *Why* do they feel this?, we should ask. Certainly it is *possible* that the feelings are meaningless epiphenomena of depression – even though feelings are not normally meaningless – but it is at least equally possible (and in our view more so) that they are *not* meaningless.

We think the most obvious way of making sense of this complex of feelings is suggested by what the DSM IV definition of major depression describes as diagnostic criterion E:

‘The symptoms are not better accounted for by bereavement’ (emphasis added).

This differential diagnostic criterion suggests that depression may be easily mistaken for bereavement, which in turn suggests that depression is characterized by a complex of feelings that closely resembles those associated with grief. It therefore seems reasonable to infer that depression might have something to do with social *loss*. This reminds us of what the early psychological investigators of depression (who were not embarrassed by feelings and their meanings) concluded on the basis of *talking* to patients about what their feelings might mean to them personally: they concluded that depression was akin to grief, that it seemed to be a pathological form of mourning (Freud, 1917).

It is in fact well established today that early separation experiences do indeed predispose to depression (Heim and Nemeroff, 1999; Pryce et al., 2005), possibly through mediation of the stress cascades and possibly also via other ‘general sickness’ mechanisms (McEwen, 2007). We also know that a first depressive episode is most likely to be triggered by social loss (see from Bowlby, 1980 to Watt and Panksepp, 2009).

Affective neuroscience

In light of such commonplace observations to the effect that depressive feelings are connected with the psychology of attachment and loss, it is striking that we are *not* focusing our attention on the mammalian brain systems that evolved specifically for the purpose of mediating attachment and loss, and which produce the particular type of pain associated with these biological phenomena of universal significance, namely *separation distress* (also known as ‘protest’ or ‘panic’) which, if it does not result in reunion, is typically followed by hopeless ‘despair’.

It is likely that a mammalian brain system evolved specifically to generate these depression-like feelings (Panksepp, 1998; 2003; 2005). This brain system evolved from general pain mechanisms, millions of years ago, apparently for the purpose of forging long-term *attachments* between mothers and their offspring, between sexual mates, and ultimately between social groups in general. When such social bonds are broken through separation or loss of a loved one, or the like, then these brain mechanisms make the sufferer feel bad in a particular way. This special type of pain is called separation distress or panic. The biological value of this type of pain is that it motivates the sufferer to avoid separation, and to seek reunion with the lost object. However, if this psychobiologically desirable outcome fails to materialize, then a second mechanism kicks in, which shuts down the distress and causes the lost individual to *give up*. This giving up is the ‘despair’ phase of social loss (Panksepp et al., 1989; 1991; Watt and Panksepp, 2009).

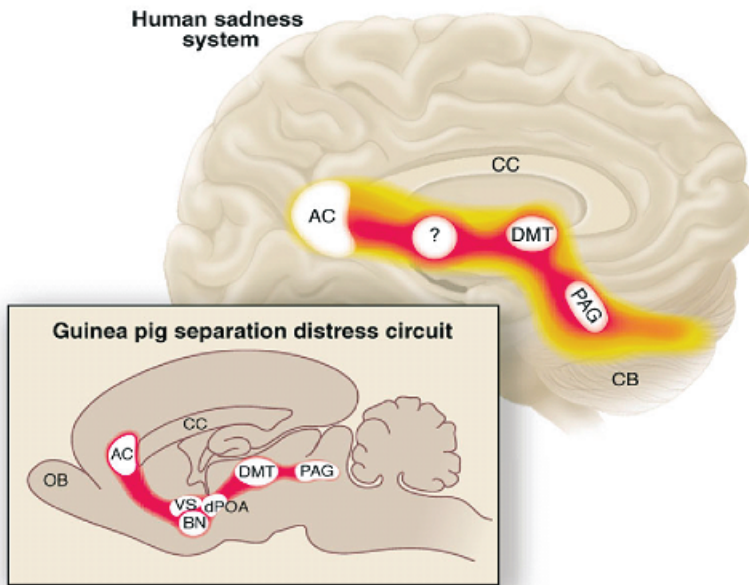


Figure 1. (From Panksepp, 2003).

This system is embodied in a well-defined network of brain structures (Figure 1), starting in the anterior cingulate gyrus (about which so much has been said in recent neuro-imaging studies and DBS treatments of depression; Mayberg et al., 2005), coursing downwards through various medial thalamic, and various basal forebrain nuclei, terminating in the ancient midbrain (pain generating) neurons of the periaqueductal grey. Regulation of this system is fundamentally mediated by opioid receptors. Mu-opioid agonists in particular activate it in such a way as to generate feelings of secure well-being that are the very opposite of depression, whereas mu opioid blockade or withdrawal promotes separation distress. This state is most readily identified in animal models by distress vocalizations (Panksepp, 1998). Bowlby (1980) classically described this phenotype as 'protest' behavior, which he contrasted with the more chronic 'despair' behaviours that immediately follow on from it. The transition from acute 'protest' to chronic 'despair' presumably evolved to protect the separated animal from metabolic exhaustion, or alternatively to deflect the attention of predators, or both.

It is the 'despair' phenotype that seems most closely to resemble clinical depression (Harris, 1989).

The separation distress system, which is regulated by the hormonal and peptide releases that precede childbirth and facilitate maternal care (e.g., oxytocin,

prolactin), developed early in mammalian evolution. This is why the mechanisms which mediate attachment and separation are much more sensitive in females – who are twice as likely as males to suffer from depression. We have also known for a long time that the chemicals that mediate the brain's separation/attachment mechanism (opioids) have powerful anti-depressant properties (Bodkin et al., 1996). If it were not for the addictive risks of opiates, they would almost certainly have formed the front line of anti-depression medications. In fact, there is good reason to believe that the natural brain chemicals – endorphins – that make us feel good when we are safely and securely attached are themselves addictive; in short, that affectionate bonds are a primal form of addiction. This system apparently provides the elemental means by which mother and infant attach to each other – the means by which they become addicted to one another.

Although these opioid-driven attachment systems may be the pivotal mechanism in depression, there may of course also be many associated mechanisms that mediate the various depressive subtypes. For example, dynorphin-facilitated shutdown of dopamine-driven appetitive systems (when an individual 'gives up' in despair) may form an independent aetiological mechanism in a subset of cases (McLaughlin et al., 2006; Nestler and Carlezon, 2006).

It seems that the pain of social loss and defeat are the price that we mammals had to pay for the evolutionary advantages bestowed by this system, that is, by mammalian social attachment, the prototype of which is the mother-infant bond. This is an instance of a more general principle: conscious feelings, both positive and negative, evolved because *they enhance survival and reproductive success*. This is their causal role.

Conclusion

So why does depression feel bad? It feels bad, on our hypothesis (initially advanced by Watts and Panksepp, 2009), for two reasons: firstly, to encourage us to form attachments, particularly to early care-giving figures, but also with our sexual mates and offspring and social groups and the like; and secondly, to persuade us to give up hope if our attempts to re-unite with such figures or groups do not succeed within a limited time-frame, when we have become detached (or lost). The fact that such feelings can be too easily provoked, or too difficult to erase, etc., in some individuals, is immaterial to the biological forces that selected them into the mammalian genome in the first place.

In light of the existence of brain structures that generate such feelings, it seems reasonable to hypothesize that the linchpin of at least one major form of

depression is none of the things that have so preoccupied contemporary psychiatric researchers over the past three decades, but rather the evolutionarily-conserved brain state that mediates the transition from 'protest' to 'despair' in the wake of social loss. In other words, it seems reasonable to hypothesize that the core brain basis of depression revolves around the process by which separation distress is normally shut down (possibly by kappa-opioids), prompting the animal to 'give up'.

Why aren't psychiatric researchers investigating the role of these candidate brain processes in depression? They seem to be the obvious place to start, if we are going to take the phenomenology of depression itself (as opposed to things that correlate with it) as our starting point.

We believe that such obvious starting points are neglected due to an ongoing, deep prejudice against acknowledging the implications for science of the subjective nature of consciousness, and its causal efficacy in the brain. This prejudice is most unfortunate, because subjective consciousness certainly exists, and it almost certainly evolved for a reason (it almost certainly enhanced reproductive fitness). It is accordingly almost certainly a central feature of how the brain works. We therefore ignore it at our peril.

References

- Berger, M., Gray, J. A. and Roth, B. L. (2009). The expanded biology of serotonin. *Annual Review of Medicine*, 60, 355–366.
- Bodkin, J. A., Zornberg, G. L., Lukas, S. E. and Cole, J. O. (1995). Buprenorphine treatment of refractory depression. *Journal of Clinical Psychopharmacology*, 15, 49–57.
- Bowlby, J. (1980). *Loss: Sadness and depression*. New York: Basic Books.
- De Kloet, E., Joels, M. and Holsboer, F. (2005) Stress and the brain: From adaptation to disease. *Nature Reviews Neuroscience*, 6, 463–475.
- Delgado, P. L., Charney, D. S., Price, L. H., Aghajanian, G. K., Landis, H. and Heninger, G. R. (1990). Serotonin function and the mechanism of antidepressant action: Reversal of antidepressant-induced remission by rapid depletion of plasma tryptophan. *Archives General Psychiatry*, 47, 411–418.
- Freud, S. (1917). Mourning and melancholia. *Standard Edition*, 14, pp. 239–258. London: Hogarth Press.
- Freud, S. (1940). An outline of psychoanalysis. *Standard Edition*, 23, pp. 144–207. London: Hogarth Press.
- Harris, J. C. (1989). Experimental animal modeling of depression and anxiety. *Psychiatric Clinics of North America*, 18, 815–836.
- Harro, J. and Oreland, L. (2001). Depression as a spreading adjustment disorder of monoaminergic neurons: A case for primary implications of the locus coeruleus, *Brain Research Reviews*, 38, 79–128.

- Heim, C. and Nemeroff, C. (1999). The impact of early adverse experiences on brain systems involved in the pathophysiology of anxiety and affective disorders. *Biological Psychiatry*, 46, 1509–1522.
- Koziek, M., Middlemas, D. and Bylund, D. (2008). Brain-derived neurotrophic factor and its receptor tropomyosin-related kinase B in the mechanism of action of antidepressant therapies. *Pharmacology and Therapeutics*, 117, 30–51.
- Levinson, D. F. (2006). The genetics of depression: A review. *Biological Psychiatry* 60, 84–92.
- Mayberg, H., Lozano, A., Voon, V., McNeely, H., Seminowicz, D., Hamani, C., Schwab, J. and Kennedy, S. (2005). Deep brain stimulation for treatment-resistant depression. *Neuron*, 45, 651–660.
- McEwen, B. S. (2007). Physiology and neurobiology of stress and adaptation: Central role of the brain. *Physiological Reviews*, 87, 873–904.
- McLaughlin, J. P., Li, S., Valdez, J., Chavkin, T. A. and Chavkin, C. (2006). Social defeat stress-induced behavioral responses are mediated by the endogenous kappa opioid system. *Neuropsychopharmacology*, 31, 1241–1248.
- Miller, A. H., Maletic, V. and Raison, C. L. (2009). Inflammation and its discontents: The role of cytokines in the pathophysiology of major depression. *Biological Psychiatry*, 65, 732–741.
- Nestler, E. J. and Carlezon, W. A. Jr. (2006). The mesolimbic dopamine reward circuit in depression. *Biological Psychiatry*, 59, 1151–1159.
- Panksepp, J. (1998). *Affective Neuroscience: The Foundations of Human and Animal Emotion*. New York: Oxford University Press.
- Panksepp, J. (2003). Can anthropomorphic analyses of “separation cries” in other animals inform us about the emotional nature of social loss in humans? *Psychological Reviews*, 110, 376–388.
- Panksepp, J. (2003). Feeling the pain of social loss. *Science*, 302, 237–239.
- Panksepp, J. (2005). Feelings of social loss: The evolution of pain and the ache of a broken heart. In R. Ellis and N. Newton (Eds). *Consciousness and Emotions*, Vol. 1. (pp. 23–55). John Benjamins: Amsterdam.
- Panksepp, J., Lensing, P. and Bernatzky, G. (1989). Delta and kappa opiate receptor control of separation distress. *Neuroscience Abstracts*, 15, 845.
- Panksepp, J., Yates, G., Ikemoto, I. and Nelson, E. (1991). Simple ethological models of depression: Social-isolation induced “despair” in chicks and mice. In B. Olivier and J. Moss (Eds.), *Animal Models in Psychopharmacology* (pp. 161–181). Holland: Duphar.
- Pryce, C. R., Ruedi-Bettschen, D., Dettling, A. C., Weston, A., Russig, H., Ferger, B. and Feldon, J. (2005). Long-term effects of early-life environmental manipulations in rodents and primates: Potential animal models in depression research. *Neuroscience and Biobehavioral Reviews*, 29, 649–674.
- Schildkraut, J. (1965). The catecholamine hypothesis of affective disorders: A review of supportive evidence. *American Journal of Psychiatry*, 122, 509–522.
- Watts, D. and Panksepp, J. (2009). Depression: An evolutionarily conserved mechanism to terminate separation distress? A review of aminergic, peptidergic, and neural network perspectives. (with commentaries). *Neuropsychanalysis*, 11, 100–104.
- Zupancic, M. and Guilleminault, C. (2006). Agomelatine: A preliminary review of a new antidepressant. *CNS Drugs*, 20, 981–992.

Dementia and the boundary between conscious and nonconscious awareness

Daniel Collerton

Dementing illnesses massively damage the structures and processes thought to underlie consciousness. However, maintained alertness, preserved awareness, and the appearance of hallucinations in people with dementia all suggest that consciousness survives extreme challenge; implying that conscious processes are robust, and conscious content invariably present.

Introduction

Last scene of all,
That ends this strange eventful history,
Is second childishness and mere oblivion,
Sans teeth, sans eyes, sans taste, sans everything.

William Shakespeare, *As You Like It*,
Act II, Scene VII, 163–166

Dementia is a human state defined by combined impairments in many aspects of cognition and accompanied by extensive emotional and psychological changes. It results from damage to the cerebral cortex by one or more of a number of different illnesses: most commonly, Alzheimer's disease, cerebrovascular disease, fronto-temporal dementia, and dementia with Lewy bodies (Ames et al., 2005). Its combined effects provide a test bed for examining the relationship between consciousness and the cognitive architecture which is thought to support it. Evidence from the effects of simultaneous impairments should compliment data from specific cognitive impairments which suggests that one aspect of consciousness can be selectively lost without consciousness as a whole failing; for example, restricted spatial neglect. However, if all cognition is compromised, is consciousness then lost? Put another way, if the structures and processes which underlie the ability to have conscious content are profoundly damaged, what happens to that content?

Do conscious processes become non conscious? The very term de-mentia implies loss of mind, and the traditional view of people with dementia has considered them as having no conscious life: the Shakespearian view – “mere oblivion ... sans everything”.

In my clinical work with people with dementia, consciousness is a concept that makes only sporadic appearance. Other concepts; mood, memory, attention – in fact, all of the other cognitive functions – are in daily use, but consciousness is used only to refer to its absence – unconsciousness – and that only rarely.

At least two explanations for this clinical invisibility come to mind. The first is that consciousness does not exist. We do not talk of it because it is not there. The second, more interesting, possibility is that consciousness is so resilient that it survives even the extreme distributed cerebral damage that dementing illnesses produce. It is so rarely absent, that its presence is unremarkable. This is the proposal that I will explore in this chapter.

Mechanisms of consciousness and the relevance of dementia

I will use consciousness to mean subjective awareness of, for example, a perception, a memory, or an emotion i.e. “the component of waking awareness perceptible by an individual at a given instant” (Grove, 2008). Non-conscious processes are thus defined by exclusion.

A number of difficulties arise immediately. The highly labile content of consciousness could blur what may be a simple division of conscious from non conscious processes into an apparently less clear cut distinction. Partial awareness of perceptions is common; for example, in the perception of visual scenes in which there is a graduation of awareness of different components of the apparently whole percept (Hayhoe and Ballard, 2005). A further problem arises in that the consciousness of one person can not be directly experienced by another, merely inferred from observation and reflection. Accurate report of cognitive content, of perceptions or memories, is the standard operational index of consciousness (Seth et al., 2005). However, the cognitive impairments of dementia challenge this yardstick. Inaccurate report in dementia does not necessarily mean a lack of consciousness; it may be the accurate report of incorrect content, or the inaccurate report of correct content.

These points argue for a broad brush, if indeed not crude, approach to the question of the effects of dementia on consciousness. Figure 1A gives a simple illustration which relates conscious cognitive content to cognitive processes. The bulk of cognitive processes are, and always will be, non conscious. A limited number of nearly conscious processes are maintained so as to be readily accessible for

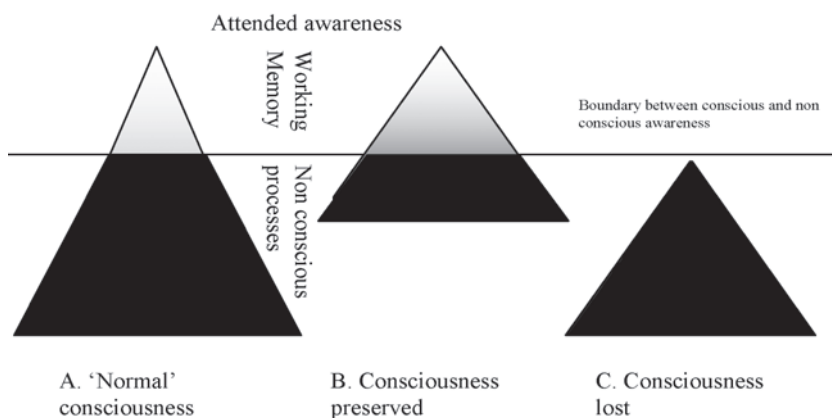


Figure 1. View of the relationship between subjective conscious awareness, and the conscious and non conscious cognitive processes which underlies it. (1A) demonstrates how most processes are non conscious with a gradation in awareness within working memory to the only fully conscious content; that which is attended to. As processes are diminished through illness, consciousness can either be preserved (1B) or lost (1C).

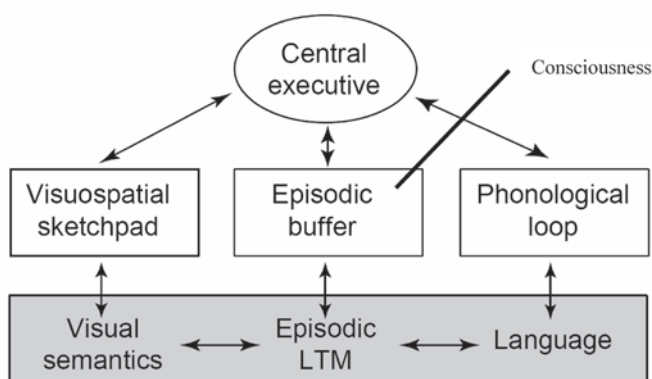


Figure 2. Processes of working memory and their relationships with non conscious processes (in gray box) from Baddeley (2003).

the task in hand and form the subsystems of working memory (Figure 2; Baars and Franklin, 2003; Baddeley, 2003). Baddeley suggests that the episodic buffer (the link between long term memory and current experience) is particularly accessible to conscious awareness. Attentional focus shifts a very limited amount of information into the episodic buffer and hence into conscious awareness. Thus,

the short term stores and systems to manipulate them which make up working memory allow the persistence of sensations as attention shifts so that a coherent consciousness can be maintained (Baddeley, 2003).

In considering the effects of damage to these processes, an appropriately simple metaphor is the iceberg. The portion that floats above the water is conscious; the portion below is non conscious. What happens as dementia melts the metaphorical iceberg? Keeping with Shakespeare, does it slip “full fathom five” below the water so that conscious awareness is lost (C), or do aspects continue to be visible (B)?

Dementia is associated with a unique combination of damage to these critical attentional and working memory systems. Studies which relate the effects of dementia specifically to consciousness (for example, Gil et al., 2001; Gregory et al., 2002; Salmon et al., 2005) are rare. In agreement with the effects of specific cognitive impairments on consciousness, they do show that relatively mild dementia has only limited effects on only some aspects of consciousness. However, they leave open the question of whether the more severe and more widespread impairments seen in advanced dementia might have greater effects. Since severe dementia precludes investigational techniques that need the understanding and co-operation of the experimental participant, a more indirect approach is needed, which separately reviews the effects of dementia on attention and working memory and relates these to consciousness; problematic though it may be to attempt to relate one type of measure (objective performance on a test of attention or memory) to another (subjective judgement of a person as having consciousness) across different areas of enquiry.

Attention, dementia, and consciousness

Attention, the ability to focus consciousness on a limited aspect of a phenomenon is greatly disturbed in dementing illnesses in all its aspects – sustaining, selecting, switching, and dividing. The effects of these impairments, however, appear not to observably reduce consciousness, but to shift it onto other subjects; possibly by tilting the balance between external stimuli and internal representations. The lack of a memory leads to the intrusion of what can be remembered, hence people misidentify their children as their parents. The lack of the explanation that one has not remembered where an item is, leads to another explanation; that it has been moved by someone else. The lack of a correct perception, leads to a hallucination. Consistent with “winner takes all” models of consciousness (Block, 2005), what people can still be conscious of takes the place of what they are no longer aware of.

This suggests that loss of the ability to attend does not abolish consciousness but instead changes its content. Some content which would not be conscious becomes so, suggesting that there remains a boundary between conscious and non-conscious processes, but that it becomes more porous. Consideration of attention also allows us to investigate the potential significance of severity of impairments. It might be that a more severe impairment of attention would destroy consciousness, or at least fragment it to the point where it is no longer functional. However, the disorder with the greatest disturbance of attention, dementia with Lewy bodies, is also the disorder with the greatest frequency of hallucinations (Collerton et al., 2005), and of fluctuations in awareness (Walker et al., 2000). Both these papers have argued that attention and these symptoms are causally related. Though this does not rule out the possibility of a threshold of attentional impairment beyond which consciousness does disintegrate, it does argue against a simple linear relationship. That increasingly poor attention does not lead to loss of consciousness but to increasing fluctuations, suggests that there are robust recovery mechanisms that can maintain conscious processes when they are disturbed.

This suggests that even when the ability to maintain a conscious focus on veridical perceptions or memories is lost, consciousness is maintained – there is always something in mind. To return to the iceberg metaphor; if one part rotates below the waves, another rises to take its place. It may occasionally sink beneath the waves, but it bobs again to the surface.

Working memory, dementia and consciousness

If the loss of the ability to attend does not lead to the fragmentation of consciousness, what if the psychological processes – working memory – which underlie conscious awareness are destroyed? It is well known that damage to one aspect of working memory, the phonological loop as an instance (Vallar and Papagno, 2002) does not lead to generally disturbed consciousness. Similarly, anosognosia leads to a loss of consciousness awareness for only some aspects of the self (Orfei et al., 2007). Anosognosia suggests that a loss of awareness of one concept does not lead to a diminution of consciousness as a whole; rather that consciousness expands to fill the gap, since otherwise the person would be aware of the impairment.

However, what happens if several systems are disturbed simultaneously as in Alzheimer's disease or dementia with Lewy bodies (Salmon et al., 2008)? Do people with dementia have a recognisable, human consciousness, despite these severe impairments in working memory tasks?

In contrast with the Shakespearean concept of dementia, more recent views have emphasized the preserved basic humanity of people with even severe

dementia (Kitwood, 1997). In the new conceptualisation which has developed from Kitwood and others' work, the consciousness of the person with dementia is not primarily of the dementia itself, but of a person who is actively trying to cope with the effects of dementia. Hence, people with dementia continue to think, act, and make sense of the world as best they can; implying, or at least assuming, that consciousness continues.

Thus, as with attention, reduced working memory seems to manifest itself not as a loss of consciousness, but as a restriction in its capacity. What may be lost is not awareness *per se*, but the ability to stand at one remove from it – to be aware of being aware (Gil et al., 2001; Salmon et al., 2005).

Cortical coherence in dementia

If awareness is maintained, is there evidence that despite massive cortical cerebral damage (the cortex can lose up to 50% of its volume in severe dementia, Esiri and Morris, 1997), the brain processes underlying consciousness continue to function?

Cortical coherence, indexed by gamma oscillations, has been suggested as a measure of the cortical binding thought to be necessary for an integrated conscious awareness (Ortinski and Meador, 2004; Rees et al., 2002). In dementing illnesses, damage to the brain structures (primarily thalamo-cortical) essential to coherence is common, and gradually increases as the disease progresses. Though there is relatively little data, measures of cortical coherence suggest that coordinated brain activity in the gamma range is robust in dementia (van Deursen et al., 2008) while activity in other bands declines (Czigler et al., 2008; Parka et al., 2008; Rossinia et al., 2007; Vesna, 2005).

Conclusion

Consciousness, at least as defined here, is robust, resilient, and always full. Though dementing illnesses greatly damage the systems which underlie awareness, consciousness appears to be remarkably unaffected. As with the consciousness of a child, consciousness in dementia is recognisably human. It may be filled with content that is idiosyncratic, it may go for a time, but it returns, and it is never empty. The boundaries between systems are maintained, and remain broadly functional. Just as in sleep, what may seem to an observer to be an empty state is actually filled with a glorious variety of conscious experience – or, to put it another way, consciousness abhors a vacuum.

References

- Ames, D., Burns, A. and O'Brien, J. (2005). *Dementia*. (3rd ed.). London: Hodder Arnold.
- Baars, B. J. and Franklin S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Sciences*, 7, 166–172.
- Baddeley, A. (2003). Working memory, looking back and looking forward. *Nature Reviews Neuroscience*, 4, 829–839.
- Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences*, 9, 46–52.
- Collerton, D., Perry, E. and McKeith, I. (2005). Why people see things that are not there. A novel Perception and Attention Deficit model for recurrent complex visual hallucinations. *Behavioral and Brain Sciences*, 28, 434–456.
- Czigler, B., Csikósa, D., Hidasib, Z., Gaála, Z. A., Csibrib, E., Kiss, E., Salacz, P. and Molnára, M. (2008). Quantitative EEG in early Alzheimer's disease patients – Power spectrum and complexity features. *International Journal of Psychophysiology*, 68, 75–80.
- Esiri, M. M. and Morris, J. H. (1997). *The neuropathology of dementia*. Cambridge: Cambridge University Press.
- Gil, R., Arroyo-Anllo, E. M., Ingrand, P., Gil, M., Neau, J. P., Ornon, C. and Bonnaud, V. (2001). Self-consciousness and Alzheimer's disease. *Acta Neurologica Scandinavica*, 104, 296–300.
- Gregory, C., Lough, S., Stone, V., Erzincliglu, S., Martin, L., Baron-Cohen, S. and Hodges, J. R. (2002). Theory of mind in patients with frontal variant frontotemporal dementia and Alzheimer's disease, theoretical and practical implications. *Brain*, 125, 752–764.
- Grove, P. B. (Ed.). (2008). *Webster's Third New International Dictionary*. Springfield, MA: Merriam-Webster.
- Hayhoe, M. and Ballard, D. (2005). Eye movements in natural behaviour. *Trends in Cognitive Sciences*, 9, 188–193.
- Kitwood, T. (1997). *Dementia reconsidered, the person comes first*. (Rethinking aging). Milton Keynes: Open University Press.
- Meador, K. J., Ray, P. G., Echaz, J. R., Loring, D. W. and Vachtsevanos, G. J. (2002). Gamma coherence and conscious perception. *Neurology*, 59, 847–854.
- Orfei, M. D., Robinson, R. G., Prigatano, G. P., Starkstein, S., Rusch, N., Bria, P., Caltagirone, C. and Spalletta, G. (2007). Anosognosia for hemiplegia after stroke, a systematic review of the literature. *Brain*, 130, 3075–3090.
- Ortinski, P. and Meador, K. J. (2004). Neuronal mechanisms of conscious awareness. *Archives of Neurology*, 61, 1017–1020.
- Parka, Y.-M., Cheb, H.-J., Imc, C.-H., Junga, H.-T., Baea, S.-M. and Lee, S.-H. (2008). Decreased EEG synchronization and its correlation with symptom severity in Alzheimer's disease. *Neuroscience Research*, 62, 2112–117.
- Rees, G., Kreiman, G. and Koch, C. (2002). Neural correlates of consciousness in humans. *Nature Reviews Neuroscience*, 3, 261–270.
- Rossini, P. M., Rossid, S., Babilonic, C. and Polichg, J. (2007). Clinical neurophysiology of aging brain. From normal aging to neurodegeneration. *Progress in Neurobiology*, 83, 375–400.
- Salmon, D. P., Hamilton, J. M., Filoteo, J. V., Heindel, W. C. and Galasko, D. (2008). Working memory deficits are greater in patients with autopsy-confirmed dementia with Lewy

- bodies (DLB) than in those with Alzheimer's disease (AD). *Alzheimer's and Dementia*, 4 (Supplement 1), T434–T435.
- Salmon, E., Ruby, P., Perani, D., Kalbe, E., Laureys, S., Adam, S. and Collette, F. (2005). Two aspects of impaired consciousness in Alzheimer's disease. *Progress in Brain Research*, 150, 287–98.
- Seth, A. K., Baars, B. J. and Edelman, D. B. (2005). Criteria for consciousness in humans and other mammals. *Consciousness and Cognition*, 14, 119–139.
- Vallar, G. and Papagno, C. (2002). Neuropsychological impairments of verbal short-term memory. In A. Baddeley, B. Wilson and M. Kopelman (Eds.), *Handbook of memory disorders* (2nd ed.) (pp. 249–270). Chichester: Wiley.
- van Deursen, J. A., Vuurman, E. F. P. M., Verhey, F. R. J., van Kranen-Mastenbroek, V. H. J. M. and Reidel, W. J. (2008). Increased EEG gamma band activity in Alzheimer's disease and mild cognitive impairment. *Journal of Neural Transmission*, 115, 1301–1311.
- Vesna, J. (2005). EEG in dementia. Review of the past, view into the future. In S. N. Sarbahikari (Ed.). *Depression and dementia, progress in brain research, clinical applications and future trends* (pp. 245–304). Hauppauge, NY: Nova Science Publishers.
- Walker, M. P., Ayre, G. A., Cummings, J. L., Wesnes, K., McKeith, I. G., O'Brien, J. T. and Ballard, C. G. (2000). Quantifying fluctuation in dementia with Lewy bodies, Alzheimer's disease, and vascular dementia. *Neurology*, 54, 1616–1625.

Consciousness as the spin-off and schizophrenia as the price of language

Timothy J. Crow

Hearing “voices”, the experience of thoughts as not one’s own, and incoherent speech are symptoms of mental illness (usually termed “schizophrenia”). Here I argue that (1) the “torque” is the feature that defines the human brain as four chambered by comparison with the two chambers of the generalized mammalian brain, (2) by separating “thought” from speech production in the frontal lobes, and “meaning” from speech perception in occipito-parieto-temporal association cortex the torque thereby confers on the species the capacity for language, and (3) the phenomena of psychosis can be seen as “leakage” from one to another of the four quadrants of association cortex. Thus consciousness is the spin-off and schizophrenia is the price that *Homo sapiens* pays for language.

Schizophrenia as a disorder of consciousness?

In all populations a fraction of individuals from the second to fifth decades of life begins to hear voices when there is no-one to account for that experience, or becomes convinced that a conspiracy is afoot when none such exists. Others develop swings of mood into depression on the one hand and elation on the other. Such states are regarded as illnesses, schizophrenia in the first case, manic-depressive disorder in the second. But neither schizophrenia nor manic-depressive disorder are categorical disease entities but should be considered in dimensional terms with some continuity (“schizo-affective” psychoses) between them. The implication of a dimensional concept is that these factors somehow extend into the normal population. Kretschmer (1925, pp. 118–119) wrote:

“...We can never do justice to the endogenous psychoses so long as we regard them as isolated unities of disease, having taken them out of their natural heredity environment, and forced them into the limits of a clinical system. Viewed in a large biological framework, however, the endogenous psychoses are nothing other than marked accentuations of normal types of temperament”.

The problem was put in a somewhat different perspective by the evolutionary theorists Julian Huxley and Ernst Mayr. If the variation is genetic and associated with a biological disadvantage (fecundity is reduced by 30% in females and 70% in males suffering from schizophrenia) why does it persist? Moreover if the variation is present in all populations it seems its origin must be ancient and related to the origin of the species (Crow, 2000).

Psychotic symptoms have been described as disorders of consciousness (Frith, 1979; Crow, 1979). For example it has been suggested is that there is some sort of enlargement of consciousness in psychosis such that neural phenomena which are normally in the subconscious, intrude into the waking state. The question arises therefore whether the symptoms of psychosis cast light on the nature of consciousness? One can ask whether such symptoms are better described as disorders of consciousness or language?

Psychotic symptoms in the split brain: Disorders of consciousness?

Some analogues of psychotic symptoms (such as hearing voices, having thoughts inserted into ones mind, and being under the control of alien forces) have sometimes been described when input to the two hemispheres is separated in subjects who have undergone section of the corpus callosum (Gazzaniga, 1970; Nasrallah, 1985) for intractable epilepsy.

In their studies of the effects of callosal section Seymour et al. (1994) reported the anomalies as disorders of consciousness, even a separation of two consciousnesses within a single individual. According to Seymour et al. (1994) "The classic example of the disconnection syndrome is the patient whose speaking left hemisphere cannot name or identify an object held in his left hand, but who can select the object under guidance of his right hemisphere". Evidence of this sort led to the view that "the split brain patient possessed two separate minds operating in parallel".

The structure of language and its decomposition in psychosis

The nuclear (core) symptoms of schizophrenia according to Kurt Schneider include (1) voices that repeat or echo the individual's thought or comment on his/her actions, and (2) the experience that alien thoughts are inserted into, or that his/her own thoughts are withdrawn from, his/her mind:

Schneider (1959) considered that when present these symptoms identify illnesses to which we will agree to attach the label schizophrenia. In other words

he proposed an operational definition – this he suggested was the core, but he did not exclude that illnesses without these symptoms might also be a part of the same entity. That they do aggregate together and can be used to define a core syndrome (“nuclear schizophrenia”) was well established in the World Health Organization Ten-Country study of the Incidence and Manifestations of Schizophrenia (Jablensky et al., 1992).

How are the symptoms to be explained? The theory to be outlined is that the asymmetry of the human brain (the torque) is the foundation of the faculty of language – and that these are primary disorders of the structure of language, that they reveal its constituent elements and the way in which these elements are segregated within the four quadrants of heteromodal association cortex

a. *The linguistic sign is bi-hemispheric.* De Saussure (1916) maintained that the linguistic sign (the word) was characteristically bi-partite, comprising a signal (the sound pattern or phonological engram) and a signalled (the associated concept or meanings). The association between the sound pattern and its meanings, according to de Saussure, is arbitrary – any sound pattern can be associated with any concept or meaning (the first principle). The contrast resembles Chomsky’s distinction between phonemic (PF) and logical (LF) form as components of the *structural derivation* in the minimalist theory.

In de Saussure’s system there is a two-way relationship between the components – with movement from the sound pattern to the meanings in speech reception, and from the concepts to sound patterns in speech production. What is the neural basis of the separation of the two components? If asymmetry is what is characteristic of the human brain it seems that there must be a relationship between specialisation of function of the hemispheres and the feature that de Saussure identifies as the key to language. A parsimonious hypothesis is that the components are (at least to a first approximation) segregated to the two hemispheres.

From Broca’s and Wernicke’s observations it is clear that what is localised in the dominant hemisphere is the phonological engram. It follows that some part of the signifieds must be assumed to be located in the non-dominant hemisphere. For each phonological engram there must be a corresponding engram – a mirror image – in the non-dominant hemisphere, but one that is systematically transformed by the differing terminations of the inter-hemispheric connexions in that hemisphere.

De Saussure’s second principle is that speech is linear – just a “ribbon of sound”. Allied to this is the notion that there must be a speaker and a hearer – speech is necessarily communicative – and the ribbon of sound is what travels between them (Figure 1).

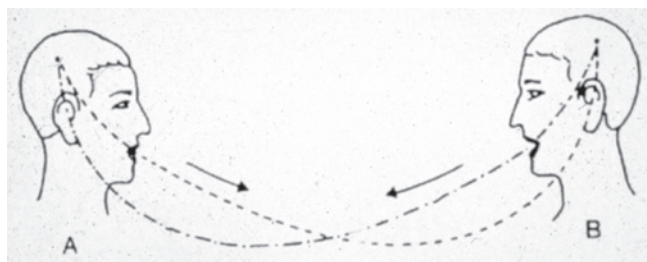


Figure 1. De Saussure's conception of the speaker and hearer and the "ribbon of sound".

One envisages therefore that speech is encoded, and this is a bi-hemispheric process, by the speaker from his concepts or thoughts into phonological engrams that are then transformed into the ribbon of sound, and that this is received by the hearer, and decoded into his own meanings or concepts, and that this encoding and decoding takes place in part by inter-hemispheric interaction. Communication depends upon the hearer sharing at least some of the speaker's signifier-signified associations, in other words that they speak the same language.

b. *Deixis and the significance of the indexical.* The system works well so long as the speaker refers to the world outside himself and the hearer. But a complication arises when he refers to himself. As Hurford (1992) points out such a referral necessitates further de-coding on the part of the hearer – that the "I" that the speaker refers to, relates not to the "I" of the hearer but to the "you" (to him) of the speaker. This class of symbols – the "indexicals" – referring to the speaker or hearer belongs to the wider class of "deictic" symbols that include reference to the "here" of the present place and the "now" of the present moment in time.

According to Buehler (1934) this triad of terms is the coordinate origin around which language is structured – this place, at the present moment in time, defined by the "I" of the speaker, without which language has no point of reference and loses its capacity to convey meaning.

The concept of the indexical provides a clue to the nature of psychotic symptoms. While there is no general misuse of the first person pronoun in individuals with psychosis those with early onset developmental disorders such as autism and Asperger's syndrome sometimes have difficulty in acquiring the distinction between the use of "I" and "you". The more general significance is that these symbols relate to what is self-generated in speech and what is other-generated – that there is a fundamental dichotomy between speech production and speech perception – and that this dichotomy can be understood in terms of the brain torque.

c. *The human brain as a four-chambered organ.* It is often overlooked that the asymmetry of the human brain is not a simple left-right difference but a deviation across the fronto-occipital axis that transforms the brain from the standard primate and vertebrate pattern of two-chambers (anterior and posterior corresponding to motor and sensory compartments) into a four-chambered organ in which motor and sensory compartments are distinguished on the left and the right sides. The torque has the effect of differentiating the two sides of the brain by influencing the relative surface area of the cerebral cortex, but it does so in different directions in the anterior (motor) and posterior (sensory) halves. It is a remarkable fact that the volumes of the two hemispheres are closely similar. What has changed relative to other primates is that the cortex is thinned or broadened on one side relative to the other and that the direction of this gradient differs along the antero-posterior axis. The effect must be assumed to be that the distribution of inter-hemispheric connexions differs on the two sides and it is this that allows the spread of neural activity to be systematically different on one side compared to the other. But the direction of the difference is opposite in the motor and sensory halves of the brain – converging from right to left anteriorly and from left to right posteriorly. The association cortex on one side is “ballooned” (thinned and broadened) relative to the homologous cortex on the other side (Harasty et al., 2003).

The human cortex is not qualitatively different from that of the chimpanzee or any other primate. If the torque is what is characteristic of the human brain these changes, and no others, are critical to the evolution of language. The inter-connexions between areas of heteromodal cortex have changed and these changes alone must account for what is characteristic of the capacity for language.

But a direct consequence of the fact that the torque constitutes a bias across the anteroposterior axis – is that the human brain is divided into 4 quadrants – the dorsolateral prefrontal cortex (DLPFC) on the right and the left side, and occipito-parieto-temporal cortex (OPTC), again on both sides. Thus the human brain is distinguished from that of all other mammals by having 4 quadrants of association cortex, right and left motor and right and left sensory. Thus functions can be distinguished in man that are not present in other mammals and this, it is argued, gives us the route to the neural basis of language. The first consequence is that transmission between areas of association cortex has directionality. If we assume that transmission is predominantly from the area of greater to lesser size then the direction is from left to right posteriorly, ie in sensory association cortex, and from right to left anteriorly in relation to motor function. Thus is identified a circuit from left to right OPTC to right DLPFC, and then to left DLPFC. This is the *sapiens*-specific speech circuit.

The key is the separation of a phonological engram from its associations. Thus if what is segregated in the dominant hemisphere is the phonological engram – a collection of simple but heavily interconnected motor sequences – this leaves open the possibility that each of these motor patterns has connexions with engrams that are systematically different (either more diffuse if motor or more restricted if sensory) in the non-dominant hemisphere.

What happens in the dominant hemisphere is of obvious interest. The sensory phonological engram is distinct from the motor engram these engrams being located in Wernicke's and Broca's areas respectively. But the form of the engram must be different – the motor engram can have relatively direct access to motor neurones and the output, but the sensory engram is one step removed from the acoustic input – word traces need to be filtered out from the totality of incoming sensory information.

Nonetheless there is a relationship between the two. Words that are heard are recognised as related if not identical to those that are spoken. The form of the relationship presumably is what is established in the course of language acquisition. Conversely some aspect of the distinction between the two is what is lost in the case of auditory hallucinations – what is clearly intrinsically generated (whether in the process of thought or in motor planning) has activated engrams that are normally accessible only to incoming acoustic stimuli.

The simplest assumption is that contiguous functions are dealt with in anatomically related areas. Thus if the output engrams are assembled in areas of association cortex that are focussed on Broca's area in the left hemisphere one must assume that the processes that are related to this assembly but differ from it in the crucial respect that they confer upon it its arbitrariness and flexibility are located in the homologous regions in the right hemisphere. In functional terms these engrams can be loosely referred to as "thought" – the precursor of, or the plans for, speech.

In the occipito-parieto-temporal junction association areas the convergence is in the opposite direction. From the phonological engrams that have been extracted from the primary acoustic signal in auditory association areas on the left there is a convergence to a smaller area of homologous cortex on the right. Thus the many-to-one transition in this case is from left to right, and the process of simplification may be identified as the distillation of "meaning" from the linear sequence of phonology on the left.

d. *The deictic origin and the performative hypothesis.* As noted above according to Buehler, language is built around a coordinate framework with an origin in the self, the present location and the present moment in time defined by the deictic symbols "I", "here" and "now". The distinction between what is self- and what is

other-generated and to put it into words clearly relates to the division between the motor and sensory association cortices and to the differentiation of the right from the left hemisphere.

The origin of the nuclear symptoms

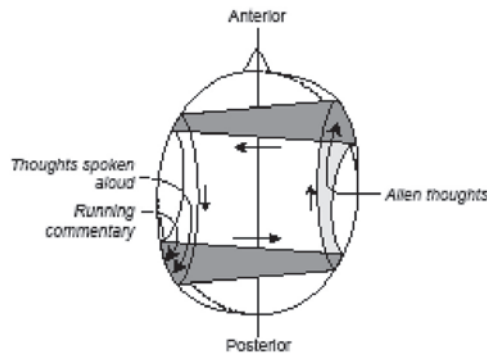


Figure 2. Schematic representation of the origin of nuclear symptoms. It is assumed that the four quadrants of hetero-modal association cortex have a degree of independence and that each has a distinct function (see also Figure 2). There are rules, presently obscure, for what is filtered by the long tracts between areas and hemispheres. The nuclear symptoms are envisaged as aberrant transmission (backflow) between areas that thus may be described as “leakage”. In the case of alien thoughts etc. leakage is from posterior (sensory) to anterior (motor), and in the case of the specific auditory hallucinations (thoughts spoken aloud) it is from anterior to posterior.

The nuclear symptoms of schizophrenia tell us what happens when the distinction breaks down. Thought, the precursor of speech, loses its characteristic independence from the outside world – thoughts are inserted into or removed from the individual’s mind – whilst retaining the features of thought they have lost the self-generativity that defines thought as a precursor of speech. The obvious interpretation is that they are influenced by activity in posterior association cortex in a way that differs from the “normal” exchange between posterior and anterior regions. Conversely auditory hallucinations such as thoughts spoken aloud or running commentary must be supposed to represent self-generated neural activity (thoughts or plans for action) that activates phonological engrams (perhaps in the superior temporal cortex on the left side) that are normally activated by speech from another individual. In each case there is a loss of the boundary in symbolic representation (words) between what is self- and what is other-generated. In each

case we can see that the boundary has something to do with what is anterior and what is posterior in association cortex. It is suggested that “counterflow” (reverse flow within the language circuit) is the basis of symptoms.

Two historical precedences of this theory should be noted. In his book *A New View of Insanity: The Duality of Mind* Arthur Wigan (1844) proposed that the two halves of the brain have independent existence and that thoughts in one half could interfere with thoughts in the other. In his book *The Origins of Consciousness in the Breakdown of the Bicomereal Mind* Julian Jaynes (1990) outlined an origin of psychotic symptoms secondary to cerebral lateralisation as a sociological phenomenon occurring relatively late in hominin evolution, ie at the time of Greek civilisation.

Developmental anomalies of anatomical asymmetry and schizophrenia

The neural basis of the torque and its relationship to pathology has been clarified in post-mortem studies (Crow et al., 1989). There are variations in anatomical asymmetry in the general population and the extremes are associated with the symptoms described as psychotic. Thus individuals with a diagnosis of schizophrenia have been shown to have:

- a. lesser or reversed asymmetries to the left of the parahippocampal, fusiform and superior temporal gyri, the latter two gyri representing heteromodal association cortex and the parahippocampal gyrus the way station from association cortex into the limbic circuit.
- b. a reduction in leftward bias of the planum temporale.
- c. reversal of the asymmetry to the left of density of pyramidal cells in dorsolateral prefrontal cortex, together with losses of size and shape asymmetries of the cell bodies (Figure 3).

In a recent meta-analysis of voxel-based morphometry (VBM) studies Honea et al. (2005) found that the most frequent deficits recorded in schizophrenia were in the medial temporal lobe on the left side and the superior temporal gyrus bilaterally. However, in terms of lateralisation the biggest deficits were in the left medial temporal lobe and in the right anterior cingulate gyrus, that is along the axis of the torque, although notably relating to limbic cortex rather than to the neocortex itself.

The anomalies of asymmetry are subtle, but arguably the hallmark. They are associated with sex differences as also are the changes in density of fibres in the corpus callosum that accompany them. Females have more fibres than males, and relative to same sex controls, female patients with schizophrenia show a decrease in density while males show an increase (Highley et al., 1999). The general thesis

that psychosis represents a deviation in the development of asymmetry and language is supported by cohort studies which indicate that pre-psychotic children are noted as mispronouncing words, and having reading difficulties (Crow et al., 1995) and for a given level of lateralisation, having fewer words than would be predicted (Leask and Crow, 2005).

What do the changes mean? If the human brain has evolved by a process of increasing hemispheric specialisation it would be expected that brain size and asymmetry are closely related and that those functions that are pivotal, eg language, will show the greatest degree of lateralisation. Such functions might also be subject to high inter-individual variability. The pivotal parameter is the age at which the brain reaches a plateau. The level and timing of this plateau are influenced by genetic or epigenetic factors. These factors must be subject to recent evolutionary pressures. Brain changes in psychosis can be viewed as an arrest of development. The brain is less lateralised and the cortical mass is smaller than might otherwise be.

Psychotic symptoms as disorders of consciousness of language?

When discussing the relationship between consciousness and language, Bickerton (1995) distinguishes three types of the former. Consciousness 1. is awareness of the environment, the ability to respond to stimuli, avoid predators etc., it interacts with other aspects of brain function such as memory and includes awareness of internal states, eg insensitivity to pain. Bickerton considers that consciousness 1. resembles what Edelman refers to as primary consciousness and that this excludes "consciousness of being conscious."

While consciousness 1. is concerned solely with online processing, in other words mediating between the nervous system and the outside world, consciousness 2. provides an awareness of consciousness 1. "We not only see the rabbit; we know that we are seeing the rabbit". Bickerton explains for consciousness of consciousness to arise there would have to be some part of the brain that is to some extent detached from the incessant traffic between the brain and environment. He likens this to Minsky's concept (1985) that Minsky describes as follows "Divide the brain into two parts (A) and (B), connect the (A) brain's inputs and outputs to the real world – so it can sense what happens there. But don't connect the (B) brain to the outer world at all. Instead connect it so the (A) brain is the (B) brain's world". Bickerton regards Minsky's (A) brain as equivalent to his primary representational system and the (B) brain to his secondary representational system which is necessary in his concept for language. However, Bickerton distinguishes a third type of consciousness which is explicitly associated with language but concedes

that there is no representative organism that has what he calls consciousness 2. but lacks language. It is a theoretical possibility only. Therefore Ockham's razor requires that the secondary representational system necessary for language applies to consciousness 2. and that is what has to be explained. Consciousness 2. in Bickerton's terms or (B) brain in Minsky's terms can be considered as equivalent to the right hemisphere, ie the part of the brain that is removed from direct interaction with the environment.

Bickerton considers that the phenomena of consciousness arises from the existence of the secondary representational system. In an interesting discussion he proposes that the invention of the self is an inevitable consequence of the secondary representational system, i.e. of consciousness 2. As I have suggested here the phenomena of psychosis can be regarded as evidence of the breakdown of the relationship between representational systems 1. and 2., and these are identified with the two sides of the brain, further divided by the motor and sensory dichotomy into the four quadrants as outlined here. The phenomena of the divided brain as described by Sperry (1968) and Gazzaniga (1992), and the phenomena of psychosis, therefore both illustrate the operation of the interaction of the two sides of the brain to yield an integrated whole of coherent language in a normal individual. Gazzaniga speaks of "the left brain interpreter" necessary for the operation of language. What the four quadrant concept adds is that one must consider language perception and production as independent interactions between the two hemispheres that operate opposing directions.

Stamenov (2003) has cast further light on these questions from a linguistic viewpoint. He writes "self-consciousness, on the other hand requires symbolic means by which the self can represent itself as a component of an explicit cognitive representation at any time as 'the same self.' Here language becomes indispensable for the development and maintenance online of self-consciousness in the individual. The foundational symbolic means for such representation in language is the first-person personal pronoun "I". This "I" is the most unusual symbol humans have at their disposal." And then later "The peculiarity of the first two aspects of "I"'s function is as follows: there is no possibility of linking on a regular basis the reference and the meaning of self without the mediation of the linguistic sign of "I". To serve its function, the "I" must refer in one direction – toward the implicit central executive, while it gets its meaning from another world, the world of current explicit cognitive representation. In order to face these two incommensurable requirements (the first negating the possibility of having representational content, the other requiring it) put by its reference and meaning, the linguistic "I" acquires the characteristics of nonspectularity – it becomes part of the explicit linguistic content without being capable of being visualized (in the way one can visualize a chair)."

“Two brain circuits provide the opportunity for monitoring one’s own speech performance which serve apparently different purposes. One monitors outer speech for errors like slips of the tongue, incorrect pronunciation or inappropriate intonation patterns. The other monitors inner speech more on the level of what one intends to say on the conceptual-intentional level (both in the case when one plans afterwards to utter behaviourally or when one thinks for himself/herself.” Again one can see these two circuits as related to the four quadrants and the two hemispheres: articulatory-perceptual and conceptual-intentional on the left.

Stamenov distinguishes between self awareness, which appears to correspond to Bickerton’s consciousness 1., and self-consciousness which approximates to Bickerton’s consciousness 2. Regarding the latter, Stamenov writes “The nonspecular self and the absent nonspecular self become possible only with the means of language. It is language that provides the symbolic means for continuous tracking of the self and empowers it with specific syntactic and semantic features. At the same time, it always keeps updated the deictic (pragmatic) her-and-now link to the executive functions of mind via pronominal reference and modality providing the self with a sense of integrity and uniqueness. Thus, language radically broadens the potential of self-awareness to a coherent and extended-in-time singular self-consciousness. It is this finely tuned cognitive orchestra that becomes shattered in schizophrenia as an illness of the self”.

If in schizophrenia there is an arrest of development (arising from diversity that is specific to the core characteristic) such that the allocation of functions to the hemispheres is prematurely curtailed, could it be that the effect is a functional separation of the two sides of the brain, ie the hemispheres acquire a degree of autonomy that is absent in normal individuals? According to this concept there is a balance between overall cerebral growth and hemispheric specialisation, and the ultimate organisation of the brain including full capacity for social communication is established only at the peak of the reproductive period of life, ie the phase of mate selection and reproduction. The most characteristic symptoms of schizophrenia may indeed be a consequence of a split brain, occurring at the asymptote of development (Crow, 1998); the components of language fall apart, and consciousness fragments.

References

- Angrilli, A., Spironelli, C., Elbert, T., Crow, T. J., Marano, G. and Stegagno, L. (2009). Schizophrenia as failure of left hemispheric dominance for the phonological component of language. *PLoS One*, 4, 1–9.
- Bickerton, D. (1995). *Language and Human Behavior*. Seattle: University of Washington.
- Buhler, K. (1934). *Sprachtheorie*. Translated (1990) by D. W. Goodwin as *Theory of Language: The Representational Function of Language*. Amsterdam: John Benjamins.
- Crow, T. J. (1979). Schizophrenia: The nature of the psychological disturbance and its possible neurochemical basis. In G. E. W. Wolstenholme & M. O'Connor (Eds.), *Brain and Mind* (pp. 335–343). Amsterdam: Excerpta Medica.
- Crow, T. J. (1998). Schizophrenia as a transcallosal misconnection syndrome. *Schizophrenia Research*, 30, 111–114.
- Crow, T. J. (2000). Schizophrenia as the price that *Homo sapiens* pays for language: A resolution of the central paradox in the origin of the species. *Brain Research Reviews*, 31, 118–129.
- Crow, T. J., Ball, J., Bloom, S. R., Brown, R., Bruton, C. J., Colter, N., Frith, C. D., Johnstone, E. C., Owens, D. G. and Roberts, G. W. (1989). Schizophrenia as an anomaly of development of cerebral asymmetry. A postmortem study and a proposal concerning the genetic basis of the disease. *Archives of General Psychiatry*, 46, 1145–1150.
- Crow, T. J., Done, D. J. and Sacker, A. (1995). Childhood precursors of psychosis as clues to its evolutionary origins. *European Archives of Psychiatry and Clinical Neuroscience*, 245, 61–69.
- Frith, C. D. (1979). Consciousness, information processing and schizophrenia. *British Journal of Psychiatry*, 134, 225–235.
- Gazzaniga, M. S. (1970). *The Bisected Brain*. New York: Appleton-Century-Crofts.
- Gazzaniga, M. S. (1992). *Nature's Mind*. New York: Basic Books.
- Harasty, J., Seldon, H. L., Chan, P., Halliday, G. and Harding, A. (2003). The left human speech-processing cortex is thinner but longer than the right. *Laterality*, 8, 247–260.
- Highley, J. R., Esiri, M. M., McDonald, B., Cortina-Borja, M., Herron, B., Cooper, S. J. and Crow, T. J. (1999). The size and fibre composition of the corpus callosum with respect to gender and schizophrenia: A post mortem study. *Brain*, 122, 99–110.
- Honea, R., Crow, T. J., Passingham, D. and Mackay, C. E. (2005). Regional deficits in brain volume in schizophrenia: a meta-analysis of voxel-based morphometry studies. *American Journal of Psychiatry*, 162, 2233–2245.
- Hurford, J. R. (1992). An approach to the phylogeny of the language faculty. In J. A. Hawkins and M. Gell-Mann (Eds.), *The Evolution of Human Languages* (pp. 273–303). Reading: Addison-Wesley.
- Jablensky, A., Sartorius, N., Ernberg, G., Anker, M., Korten, A., Cooper, J. E., Day, R. and Bertelsen, A. (1992). Schizophrenia: Manifestations, incidence and course in different cultures. A World Health Organization Ten Country Study. *Psychological Medicine*, Suppl. 20, 1–97.
- Jaynes, J. (1990). *The Origins of Consciousness in the Breakdown of the Bicameral Mind* (first published 1976). Boston: Houghton Mifflin.
- Kretschmer, E. (1925). *Physique and Character*. London: Kegan Paul, Trench, Trubner & Co Ltd.
- Leask, S. J. and Crow, T. J. (2005). Lateralization of verbal ability in pre-psychotic children. *Psychiatry Research*, 136, 35–42.

- Minsky, M. L. (1985). *The Society of Mind*. New York: Simon and Schuster.
- Nasrallah, H. A. (1985). The unintegrated right cerebral hemispheric consciousness as alien intruder: A possible mechanism for Schneiderian delusions in schizophrenia. *Comprehensive Psychiatry*, 26, 273–282.
- Saussure de, F. (1916). *Course in General Linguistics*. (Translated by R. Harris 1983 and Published by Open Court, Illinois) Paris: Payot.
- Schneider, K. (1959). *Clinical Psychopathology* (Translated by M. W. Hamilton and E. W. Anderson) New York: Grune & Stratton.
- Seymour, S. E., Reuterlorenz, P. A. and Gazzaniga, M. S. (1994). The disconnection syndrome – Basic findings reaffirmed. *Brain*, 117, 105–115.
- Sperry, R. W. (1968). *Mental Unity Following Surgical Disconnection of the Cerebral Hemispheres*. New York: Academic Press.
- Stamenov, M. I. (2003). Language and self-consciousness: Modes of self-presentation in language structure. In T. L. Kircher and A. David (Eds.), *The Self in Neuroscience and Psychiatry* (pp. 76–104). Cambridge: Cambridge University Press.
- Wigan, A. L. (1844). *A New View of Insanity: The Duality of Mind*. London: Longman, Brown, Green and Longmans.

Consciousness and psychosis associated with schizophrenia

The role of Cornu Ammonis Region 3

Ralf-Peter Behrendt

Information encoded by the autoassociation network of CA3 situates landmarks and objects within an allocentric frame of space and time. Guiding locomotion across the spatial environment, and generally organizing behaviour that transcends space and time, the hippocampus creates phenomenal space and time themselves, thus laying the foundations for conscious awareness. It is argued that conscious experience describes the informational content of self-organizing activity patterns in CA3. Excessive pyramidal cell activity in CA3, due to deficient inhibition by GABAergic basket interneurons, leading to event memory formation unrestrained by entorhinal input, may be a mechanism for the generation of hallucinations in schizophrenia.

Introduction

All conscious experience is 'situated': conscious perception and imagery are always allocentric, in that landmarks or objects are experienced within a framework of space or time. Navigation along landmarks can only be effective if representations of landmarks are encoded with reference to an allocentric spatial context that appears to be 'fixed' to the external environment. What we experience consciously is but an epiphenomenon of allocentric representations evolved for the purpose of navigation. It is suggested that situated landmarks evolved into situated objects that are linked to affective behaviour modes, whereas the spatial context in which landmarks are situated evolved into a more abstract framework of space and time. Phenomenal time and space are not dissimilar: the abstract domain of time gets its relational structure from the concrete domain of space (Boroditsky, 2000). By continuously creating allocentric representations, that is, encoding event memories, the hippocampus *creates* phenomenal space and time

themselves, consistent with philosophical insights that we do not live in an objective framework of space and time but that the spatially and temporally extending world that appears to surround us is an entirely subjective and fundamentally private creation (Behrendt, 2007).

Allocentric representations

The lateral posterior parietal cortex contains multiple representations of space (retinocentric maps and other egocentric reference frames) for the purpose of guiding particular types of motor acts (e.g., saccadic eye movements, reaching, grasping). Different ‘action-oriented spatial representations’ are implemented in different parts of the lateral posterior parietal cortex (Andersen and Buneo, 2002). In the medial parietal cortex, the precuneus and retrosplenial cortex integrate various feature representations activated in the posterior parietal cortex in order to encode, in egocentric head-centred coordinates, representations of locations or landmarks that are visible from the animal’s current position. The posterior cingulate and adjacent precuneus in the medial parietal cortex represent the visual periphery and support broad monitoring of the environment. The posterior cingulate and retrosplenial cortices are interconnected with the medial temporal lobe, particularly the parahippocampal cortex and entorhinal cortex. Sensory information processed by the posterior parietal cortex is thus transformed into egocentric head-centred coordinates (reflecting the current head direction), before being transformed into an allocentric *representation of the space out there* by the medial temporal cortex (Byrne et al., 2007). Parietal regions that form part of the ‘default-mode network’, including the posterior cingulate, precuneus, retrosplenial cortex and inferior parietal lobule, are functionally connected (showing correlated slow BOLD signal fluctuations) with the hippocampal formation (hippocampus and entorhinal cortex) (Greicius et al., 2004).

Allocentric representations of spatial locations or landmarks are encoded in the medial temporal cortex (parahippocampal regions and hippocampus). The hippocampus, in particular, ties various ‘visual feature units’ together, using ‘attractor dynamics’, to form an ‘allocentric map’ of the environment (Byrne et al., 2007). The subiculum is the main output region of the hippocampus. Its projections to the mammillary bodies forms part of Papez’s circuit. Mammillary bodies project, via the anteroventral thalamic nucleus, to the cingulate cortex, which, in turn, feeds back, via the parahippocampal gyrus, to the hippocampus. Hippocampal output via Papez’s circuit to the retrosplenial cortex may transform allocentric representations back into egocentric head-centred representations. An egocentric representation of the current head direction, encoded by the retrosplenial cortex,

then updates retinocentric and other egocentric representations in posterior parietal areas (Byrne et al., 2007).

Papez's circuit is crucial not only for locomotor navigation but also episodic memory recollection, consistent with findings that mental imagery activates medial parietal regions, but also involves medial temporal regions. The precuneus is activated during recall of visuospatial memories and imagery of outcomes of voluntary actions. The posterior cingulate cortex and precuneus are activated during vivid reexperiencing of autobiographical memories (Gilboa et al., 2004). The precuneus may encode visuospatial episodic memories by functioning as a 'connecting node' between representations of various aspects of the memory stored in posterior parietal areas. The transformation of allocentric into egocentric representations during visuospatial imagery of environmental locations would produce 'mock motor efferences'. 'Real or imagined self motion' would update egocentric representations of locations in the environment (Byrne et al., 2007). Byrne et al. (2007) suggested that an 'egocentric parietal image' is formed for 'conscious inspection' during imagery of environmental layouts or when thinking about past or future events. In contrast to position advocated by Byrne et al. (2007), mental imagery may not be egocentric or 'generated parietally', despite involvement of parietal cortices, but may be based on allocentric representations formed in medial temporal regions. Transformations from allocentric to egocentric representations would still be required for the sequential advancement of mental imagery (episodic memory recall).

Episodic memory

Layer II of the entorhinal cortex projects, via the perforant path, to the granule cell layer of the dentate gyrus. Granule cells, in turn, send mossy fibres to CA3 (and CA2). Pyramidal neurons of CA3 are extensively interconnected via recurrent axon collaterals. CA3 is thought to form a single 'autoassociation network' that displays 'attractor dynamics'. The high degree of internal connectivity and effective synaptic plasticity ('long-term potentiation') enable the rapid formation of associations among individual elements of a patterned input (Rolls, 2007). Inputs reflecting sensory detail processed in neocortical modules generate a unique identifying pattern in CA3 that acts as an 'index' linking disparate neocortical modules. The autoassociative network of CA3 stores 'arbitrary associations' between object and place information as short-term memories for durations of seconds to minutes, during which stored information may circulate through the network. The neural network of the dentate gyrus, projecting via mossy fibres to CA3, implements a 'pattern separation' process, which detects changes in

sensory input processed by the entorhinal cortex and serves the creation, in CA3, of 'orthogonal pattern representations' of the environment (Kesner, 2007; Rolls, 2007). As event memories are sequentially encoded within the autoassociative network of CA3, 'backprojections' from CA3 pyramidal cells modify synaptic connections in the dentate gyrus, with the effect of linking one event memory with the next in a sequence (Lisman et al., 2005). These arrangements may be essential for producing a spatially continuous and temporally coherent representation of the external world.

Neuronal ensembles in CA3 of the dorsal (posterior) hippocampus encode novel spatial and contextual information (Kesner, 2007). The ventral (anterior) hippocampus is preferentially involved in acquisition and retrieval of nonspatial memories. CA3 of the ventral hippocampus encodes event memories consisting of arbitrary associations between objects and places (Rolls and Xiang, 2005). Spatial information about the location of objects encoded in the parietal cortex enters the dentate-CA3 network via the medial entorhinal cortex, while information about the nature of objects provided by the temporal visual cortex enters the anterior dentate-CA3 network via the lateral entorhinal cortex (Kesner, 2007). Information about different rewards and punishers, represented in the amygdala and orbitofrontal cortex, reaches the anterior hippocampus via the lateral entorhinal cortex. Thus, CA3 of the anterior hippocampus can form arbitrary associations between information about the 'space out there' (allocentric space) and reward information received from the orbitofrontal cortex and amygdala (Rolls and Xiang, 2005). CA2 may form memories of social encounters, assimilating contextual factors, such as social odours, into representations of social events.

Changes in internal recurrent connectivity of hippocampal region CA3 encode patterned information in a manner that later enables partial fragments of a previous input pattern to retrieve the complete pattern of activity encoding the original event. Layer II of the entorhinal cortex provides direct input, via the perforant path, to CA3. Retrieval of patterns encoded in CA3 is initiated by direct entorhinal inputs representing partial retrieval cues (Kesner, 2007; Lisman et al., 2005; Rolls, 2007). Pyramidal neurons of CA3 (and CA2) project to region CA1 via Schaffer's collaterals. CA1 links several patterns of CA3 activity, representing arbitrary associations, to form an episodic or 'whole scene' memory (Rolls, 2007). CA1 receives input from layer III of the entorhinal cortex, again via the perforant path, which may support its function of representing places in a familiar environment (Kesner, 2007) (view-independent allocentric representations). CA3 is not necessary for spatial recognition memory or for the establishment and maintenance of place representations in region CA1, but is required for the recall of remote locations or trajectories toward these locations (Brun et al., 2002). Neural ensembles in CA3 that encode spatial representations of locations ahead of the

animal ('nonlocal representations') are transiently reactivated during navigational decision-making (Johnson and Redish, 2007). During episodic memory retrieval, triggered by a sensory cue, CA1 receives a time-compressed readout of a sequence of event memories encoded in CA3. Readout of sequential information through autoassociative recall provides sequential predictions to CA1. The prediction of what will happen next, based on previous experience, is compared with sensory input arriving from layer III of the entorhinal cortex.

θ Oscillations

Parvalbumin-expressing GABAergic basket cells synapse onto somata and initial axon segments of pyramidal cells in CA1, CA2, CA3 and the dentate gyrus, tightly regulating the firing activity of pyramidal cells. Fast-firing basket cells produce inhibitory postsynaptic currents in pyramidal cells, the summation of which is responsible for γ oscillations in local field potentials. During exploratory locomotion and REM sleep, neurons in superficial layers of the entorhinal cortex and their projection targets in the hippocampus discharge synchronously at γ frequencies in relation to the phase of the θ cycle. While γ oscillations in dentate gyrus and CA1 are modulated by entorhinal input, γ -oscillatory patterns in CA3 self-organize (Montgomery and Buzsáki, 2007). Basket cells, causing rhythmic inhibition in hippocampal pyramidal neurons, also play a role in the generation of θ rhythm. Synchronous γ firing in subsets of CA3 pyramidal neurons that are tuned to θ rhythm is necessary for the temporary storage of information during exploration. Cholinergic and GABAergic input from the medial septum enhances memory encoding by increasing θ oscillations in the hippocampal formation. Hippocampal θ oscillations, which are selectively present during exploratory locomotion, behavioural preparedness and attentiveness to stimuli (but also REM sleep), ensure the continuous gathering of information about the environment. If exploratory behaviour leads to contact with reward or punishment, short-term, θ -dependent memory traces are consolidated. In CA1, assemblies representing past, present and future locations are activated sequentially within a single θ cycle during spatial navigation (Shapiro and Ferbinteanu, 2006). θ Oscillations are required for coding sequential places along a track, similarly to how the formation of episodic memories, that is, learning a temporal sequence of events, involves θ oscillations. θ Periods are necessary for tying together neuronal assemblies representing sequential events whatever their nature. θ Rhythm can be found in structures around the Papez's circuit and was suggested to coordinate the flow of information between medial parietal and medial temporal regions, enabling the transformation between egocentric and allocentric representations, during

navigational tasks. θ Rhythm would also coordinate the interaction between medial temporal and medial parietal regions during recall, in imagery, of episodic memories (Byrne et al., 2007).

Seeking to explain psychophysical findings indicative of the discrete (rather than continuous) nature of conscious perception, VanRullen and Koch (2003) argued that θ oscillations, acting as a carrier for γ oscillations, provide the context for conscious memory recall, whereas α oscillations may provide the context or carrier for γ oscillations encoding perceptual experiences. Thus, “only the entire slow wave” (θ or α) and its inscribed contents “is made available at once to explicit perception” (VanRullen and Koch, 2003, p. 212). However, we need to consider, firstly, that conscious perception is indistinguishably intertwined with recognition based on memory retrieval. Secondly, what we experience as conscious memory recall or imagery is but a derivative of externalised perception; perception and imagery are, on a fundamental level, phenomenologically identical and should be attributable to the same process. Consciousness, be it perception of a seemingly external world or ‘internal’ imagery, is argued to consist of a stream of complex symbols that describe self-organizing patterns emerging in CA3. The complexity of our experience at any one time reflects the complexity of arbitrary associations formed by attractor dynamics of the autoassociation network. Thus, it is suggested that the formation of view-dependent allocentric representations in the CA3-dentate network, for the purpose of their episodic memorization, underlies conscious awareness during both perception and imagery, consistent with Zeman’s (2001) “unpalatable conclusion” that “experience is a beautiful but functionally irrelevant embellishment of physical processes” (p. 1284). Contemporary theories of consciousness tend to assume “that consciousness plays an important, more or less continuous role in directing our waking behaviour”, meaning that “consciousness makes a difference” (p. 1282). However, psychophysical evidence suggests that “consciousness arises too late to perform the kinds of function which these theories envisage”, so that it seemed possible that “rather than guiding action from moment to moment, consciousness serves its biological purpose on a longer, more reflective, time scale” (Zeman, 2001, p. 1281). Chalmers (1996) had previously argued that consciousness, being imbued with nonphysical properties, could not make a difference to the physical trajectory of behaviour. Moreover, there is a long tradition of philosophical idealism consistent with the notion that consciousness, unfolding in a realm distinct from physical reality, cannot be efficacious with regard to the physical trajectory of behaviour (‘mind-body problem’) (Behrendt, 2007). Consciousness has to be seen as forming a dimension that is orthogonal to the physical processes that underpin it; the consciously experienced world, including phenomenal space and time, occupies a dimension that is orthogonal, and thus irrelevant, to material reality.

Dreaming and REM sleep

Wakeful perception, hallucinations and dreams are phenomenologically closely related (Behrendt, 2007), and understanding the mechanism of one variety of conscious experience helps us to understand the others. Sequences of patterned ensemble firing in the hippocampus reflecting conscious experiences during wakefulness are reproduced during REM sleep. Reactivated sequences of patterned ensemble firing were suggested to represent reactivated episodic memory traces that form the content of dream states (Louie and Wilson, 2001). θ And γ oscillations are highly synchronized between dentate gyrus and CA3 during REM sleep. Dentate-CA3 synchronization may reflect the processing of event memories, whereby the dentate gyrus supports pattern separation and recombination of memory fragments (Montgomery et al., 2008).

Amygdala hyperresponsivity in social phobia and schizophrenia

Psychotic episodes in schizophrenia, which typically feature delusional beliefs about harm or danger emanating from specific individuals or social groups, are often preceded or accompanied by syndromatic or subsyndromatic social phobia characterized by intense apprehensions, generalised across all social encounters, about being criticized, negatively evaluated or rejected by others (Michail and Birchwood, 2009). In neuroimaging studies, patients with social phobia, but also individuals with excessive trait anxiety, show hyperresponsivity of the amygdala, especially on the left, to others' angry or contemptuous emotional expressions (Phan et al., 2006). The amygdala contributes to the formation of emotional memories, underlying emotionally coloured conscious experience, in several ways. Firstly, the basolateral complex of the amygdala (particularly the lateral amygdalar nucleus) modulates cortical sensory processing in association areas of the anterior lateral temporal cortex. The anterior lateral temporal cortex, which forms part of the 'ventral visual stream', forwards nonspatial visual information about objects to the perirhinal cortex, lateral entorhinal cortex and, then, anterior (ventral) hippocampus. The anterior lateral temporal cortex, which, in humans, also processes visual and nonvisual information about social and emotional cues as well as semantic components of language, acts as a 'connectional hub' to the anterior hippocampus by way of perirhinal/entorhinal cortices (Kahn et al., 2008). Secondly, glutamatergic projections from the basolateral amygdala to perirhinal and entorhinal cortices promote γ oscillations in these regions by rhythmically depolarizing pyramidal neurons, thereby enabling the integration of neocortical inputs representing different sensory modalities in emotionally arousing situations (Bauer et al., 2007). Thirdly,

the basolateral nucleus of the amygdala provides strong excitatory input directly to CA3 and CA2. Thus, the basolateral nucleus of the amygdala facilitates the formation of memories for the environmental context of aversive stimulation (contextual fear memory). In a rat model of schizophrenia, excessive glutamatergic input from the basolateral amygdala to regions CA3 and CA2 causes excitotoxic reductions in GABAergic interneuron density (especially affecting presumptive fast-firing basket cells), mirroring postmortem findings of reduced GABAergic interneuron functionality in CA3 and CA2 in schizophrenia. Reductions in GABA-mediated postsynaptic inhibitory potentials in pyramidal neurons of CA3 occurred in association with an increase in hippocampal 'long-term potentiation' (Gisabella et al., 2005). Thus, increased responsiveness of the amygdala may contribute to excessive formation of event memories relating to adverse experiences.

GABAergic hypofunction in schizophrenia

Postmortem studies of schizophrenia demonstrate GABAergic hypofunction in the hippocampus, especially in CA3 and CA2. GABAergic hypofunction predominantly affects fast-spiking interneurons (basket and chandelier cells), which contain parvalbumin and synapse onto perisomatic aspects of pyramidal cells (Lisman et al., 2008). The hippocampus of patients with schizophrenia shows profound reductions in parvalbumin immunoreactivity (Zhang and Reynolds, 2002). GABAergic neurons express glutamic acid decarboxylase (GAD), which synthesizes GABA. Postmortem analysis demonstrates downregulation of isoenzyme GAD₆₇ in CA3 and CA2 in patients with schizophrenia or bipolar affective disorder (Benes et al., 2007). GAD₆₇ is more concentrated in inhibitory neurons that fire tonically, as opposed to those that have a low basal firing rate, consistent with decreased functionality of basket interneurons in schizophrenia. Decreased functionality of fast-spiking interneurons leads to a reduction of inhibitory postsynaptic potentials in pyramidal neurons and, hence, compensatory upregulation of postsynaptic GABA_A receptors on pyramidal neurons, as can be demonstrated in CA3 and CA2 of patients with schizophrenia.

Administration of NMDA-receptor antagonists can induce psychotic experiences, such as hallucinations, in healthy subjects. NMDA-receptor hypofunction resulting from genetic abnormalities has been implicated in the pathophysiology of schizophrenia (Lisman et al., 2008). Deficient GABAergic transmission in the hippocampus may be a consequence of NMDA-receptor hypofunction. Fast-spiking interneurons use NMDA receptors to 'sense' the level of activity in surrounding pyramidal neurons (by responding to extracellular glutamate) in order to adjust the synthesis and release of GABA. Malfunctioning NMDA receptors

would be relatively insensitive to glutamate levels in the surrounding milieu, which would lead to downregulation of GABA synthesis (involving GAD₆₇) and deficient GABA release (Lisman et al., 2008). Thus, NMDA-receptor hypofunction reduces the activity of fast-spiking GABAergic interneurons in the hippocampus, resulting in disinhibition of hippocampal pyramidal cells. Accumulating evidence suggests that NMDA-receptor antagonists cause symptoms of schizophrenia by reducing the output of fast-spiking GABAergic interneurons and thereby increasing pyramidal cell activity (Lisman et al., 2008).

The decrease in GABAergic tone in the hippocampus may be an indirect consequence of elevated excitatory input from the basolateral amygdala (Gisabella et al., 2005). Increased amygdalar input may upregulate cholinergic M₁ and M₃ receptors on pyramidal cells of CA3 and CA2. The medial-septal cholinergic system exerts a facilitatory influence on pyramidal-cell excitability and synaptic transmission in the hippocampus; and upregulation of M₁ cholinergic receptors on pyramidal cells would enable cholinergic input from the medial septum to exert a strong excitatory influence on hippocampal circuitry. Enhanced firing of hippocampal pyramidal cells would subject connected GABAergic cells to excessive oxidative stress, secondarily reducing GABAergic interneuron function (Gisabella et al., 2005).

Hallucinations

Electrical stimulation of the human medial temporal lobe can elicit complex visual hallucinations and sometimes auditory hallucinations, whereby stimulations of the amygdala and hippocampus are most effective in provoking complex visual hallucinations (most commonly featuring people who speak in some cases) or vivid experiences of autobiographic memories (Vignal et al., 2007). Direct stimulation of the amygdala can elicit feelings of fear and anxiety, feelings of familiarity as well as complex hallucinations. Brain lesions in or near the amygdala are associated with schizophreniform psychoses. Although hippocampal volume is often reduced in patients with schizophrenia, metabolic activity in the hippocampus, as measured by neuroimaging, is often increased, indicating an excess of excitatory neuronal activity. Hyperactivity of the left hippocampus and parahippocampal gyrus is associated with a tendency, in these patients, to experience 'positive symptoms' (hallucinations and delusions) (Liddle et al., 2000). The actual experience of auditory (Shergill et al., 2000) or visual hallucinations is accompanied by hippocampal activation in fMRI, along with activations in higher-order neocortical sensory-processing areas (Oertel et al., 2007). These findings accord with the hypothesis that hippocampal hyperactivity might underlie hallucinations and other positive symptoms in schizophrenia (Liddle et al., 2000). Reduction of regional cerebral

metabolism in the left hippocampus that was induced by a single dose of risperidone strongly predicted the reduction in severity of 'reality distortion' (delusions and hallucinations) after 6 weeks of treatment with risperidone, "consistent with the hypothesis that reduction in aberrant hippocampal firing is a prerequisite for subsequent resolution of reality distortion" (Liddle et al., 2000, p. 406).

Patients with schizophrenia, particularly those with 'disorganized schizophrenia' (characterised by prominent hallucinations and thought disorder), are impaired in Gestalt tasks that require perceptual grouping of visual stimuli. In rats, blocking neural activity in one hippocampus induced coactivation, at γ frequencies, of pyramidal cells in the contralateral hippocampus that initially fired independently (Olypher et al., 2006). As a result, the ability of rats to subgroup distal spatial stimuli and segregate them from irrelevant local stimuli was impaired, preventing these rats from effectively avoiding regions of the room where footshock was administered. Coactivation, at γ frequencies, of initially uncoupled hippocampal pyramidal neurons amounts to failure to segregate cell assemblies encoding unrelated representations. Disorganization of spike timing between cells would provide the basis on which "a pathological steady and stable state of activity" can emerge through γ synchronisation: a 'parasitic attractor' "that does not reflect reality" but a hallucination instead (Olypher et al., 2006, p. 166).

Conclusions

Allocentric representations formed by CA3 (event memories) are to be distinguished from those formed by CA1. It is argued that CA1 rapidly encodes view-independent allocentric representations of places and, possibly, more abstract spatiotemporal 'situations', insofar as 'situations' provide a context for the selection of appropriate task modes and stimulus-response transformations. Even view-dependent allocentric representations formed by the CA3 autoassociation network can not be regarded as 'conscious representations'. CA3 allocentric representations are as unconscious as CA1 allocentric representations, although the former, unlike the latter are symbolically expressed in a stream of consciousness (owing to autoassociation processes and backprojections to the dentate gyrus). Allocentric representations refer to nonconscious processes that do affect the physical trajectory of behaviour (through hippocampal output to the medial prefrontal cortex and lateral septum), however consciousness, being a symbolic expression of view-dependent (in an abstract sense) allocentric representations, does not. Behaviour is all but unconscious; consciousness has no bearing on brain processes. Consciousness unfolds in a realm that is orthogonal to behaviour and neuronal processes, and indeed to physical reality.

Overactive hippocampal pyramidal cells are thought to contribute to psychosis by increasing the number of spontaneously active dopaminergic neurons in the ventral tegmental area and, thereby, driving a hyperdopaminergic state (Lisman et al., 2008). It is argued here that increased excitability in CA3 is directly responsible for the generation of hallucinations in psychosis (although increased dopaminergic input to CA3 may contribute to enhanced 'long-term potentiation' in CA3, perhaps via D₂ receptor activation on basket interneurons, leading to decreased GAD₆₇ expression). Hallucinations are also associated with peripheral sensory impairment and various pathologies, including schizophrenia, that affect the ability of the reticular thalamic nucleus to regulate the signal-to-noise ratio of neuronal activity in sensory thalamic nuclei (Behrendt and Young, 2004). These abnormalities would predispose to patterns of self-organizing thalamocortical activity that are underconstrained by peripheral sensory input but still capable of influencing, downstream, the formation of event memories proposed to underlie conscious experience. It could be speculated that the presence of visual hallucinations, alongside auditory hallucinations, as well as prevalent formal thought disorder (as is the case in patients with 'disorganized schizophrenia') indicates pathologically increased hippocampal activity, whereas hallucinations occurring exclusively in the auditory modality in patients with 'paranoid schizophrenia' may indicate upstream 'underconstrained' thalamocortical activity that, under conditions of abnormal reticular thalamic nucleus function, is excessively driven by attentional pressures (Behrendt and Young, 2004). Thus, hallucinations, which, too, are conscious experiences of 'something situated in a context', reflect disordered assembly formation in CA3 that results from (a) neocortical activity that is insufficiently constrained by peripheral sensory input (upstream dysfunction) or (b) GABAergic hypofunction within the hippocampus (downstream).

References

- Bauer, E. P., Paz, R. and Paré, D. (2007). Gamma oscillations coordinate amygdalo-rhinal interactions during learning. *The Journal of Neuroscience*, 27, 9369–9379.
- Behrendt, R. P. (2007). The Subjectivity of the Perceived World: Psychopathology and the Mind-Body Problem from a Perspective of Idealism. In V. W. Fallio (Ed.), *New Developments in Consciousness Research*. Hauppauge, NY: Nova Science Publishers.
- Behrendt, R. P. and Young, C. (2004). Hallucinations in schizophrenia, sensory impairment, and brain disease: A unifying model. *Behavioral and Brain Sciences*, 27, 771–787.
- Benes, F. M., Lim, B., Matzilevich, D., Walsh, J. P., Subburaju, S. and Minns, M. (2007). Regulation of the GABA cell phenotype in hippocampus of schizophrenics and bipolars. *Proceedings of the National Academy of Sciences of the USA*, 104, 10164–10169.
- Boroditsky, L. (2000). Metaphoric structuring: understanding time through spatial metaphors. *Cognition*, 75, 1–28.

- Brun, V. H., Otnass, M. K., Molden, S., Steffenach, H. A., Witter, M. P., Moser, M. B. and Moser, E. I. (2002). Place cells and place recognition maintained by direct entorhinal-hippocampal circuitry. *Science*, 296, 2243–2246.
- Byrne, P., Becker, S. and Burgess, N. (2007). Remembering the past and imagining the future: A neural model of spatial memory and imagery. *Psychological Review*, 114, 340–375.
- Chalmers, D. J. (1996). *The conscious mind*. Oxford: Oxford University Press.
- Gilboa, A., Winocur, G., Grady, C. L., Hevenor, S. J. and Moscovitch, M. (2004). Remembering our past: Functional neuroanatomy of recollection of recent and very remote personal events. *Cerebral Cortex*, 14, 1214–1225.
- Gisabella, B., Cunningham, M. G., Bolshakov, V. Y. and Benes, F. M. (2009). Amygdala-dependent regulation of electrical properties of hippocampal interneurons in a model of schizophrenia. *Biological Psychiatry*, 65, 464–472.
- Greicius, M. D., Srivastava, G., Reiss, A. L. and Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: Evidence from functional MRI. *Proceedings of the National Academy of Sciences of the USA*, 101, 4637–4642.
- Johnson, A. and Redish, A. D. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *The Journal of Neuroscience*, 27, 12176–12189.
- Kahn, I., Andrews-Hanna, J. R., Vincent, J. L., Snyder, A. Z. and Buckner, R. L. (2008). Distinct cortical anatomy linked to subregions of the medial temporal lobe revealed by intrinsic functional connectivity. *Journal of Neurophysiology*, 100, 129–139.
- Kesner, R. P. (2007). Behavioral functions of the CA3 subregion of the hippocampus. *Learning and Memory*, 14, 771–781.
- Liddle, P. F., Lane, C. J. and Ngan, E. T. (2000). Immediate effects of risperidone on cortico-striato-thalamic loops and the hippocampus. *The British Journal of Psychiatry*, 177, 402–407.
- Lisman, J. E., Talamini, L. M. and Raffone, A. (2005). Recall of memory sequences by interaction of the dentate and CA3: A revised model of the phase precession. *Neural Networks*, 18, 1191–1201.
- Lisman, J. E., Coyle, J. T., Green, R. W., Javitt, D. C., Benes, F. M., Heckers, S. and Grace, A. A. (2008). Circuit-based framework for understanding neurotransmitter and risk gene interactions in schizophrenia. *Trends in Neurosciences*, 31, 234–242.
- Louie, K. and Wilson, M. A. (2001). Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. *Neuron*, 29, 145–156.
- Michail, M. and Birchwood, M. (2009). Social anxiety disorder in first-episode psychosis: Incidence, phenomenology and relationship with paranoia. *The British Journal of Psychiatry*, 195, 234–241.
- Montgomery, S. M. and Buzsáki, G. (2007). Gamma oscillations dynamically couple hippocampal CA3 and CA1 regions during memory task performance. *Proceedings of the National Academy of Sciences of the USA*, 104, 14495–14500.
- Montgomery, S. M., Sirota, A. and Buzsáki, G. (2008). Theta and gamma coordination of hippocampal networks during waking and rapid eye movement sleep. *The Journal of Neuroscience*, 28, 6731–6741.
- Oertel, V., Rotarska-Jagiela, A., van de Ven, V. G., Haenschel, C., Maurer, K. and Linden, D. E. (2007). Visual hallucinations in schizophrenia investigated with functional magnetic resonance imaging. *Psychiatry Research*, 156, 269–273.

- Olypher, A. V., Klement, D. and Fenton, A. A. (2006). Cognitive disorganization in hippocampus: A physiological model of the disorganization in psychosis. *The Journal of Neuroscience*, 26, 158–168.
- Phan, K. L., Fitzgerald, D. A., Nathan, P. J. and Tancer, M. E. (2006). Association between amygdala hyperactivity to harsh faces and severity of social anxiety in generalized social phobia. *Biological Psychiatry*, 59, 424–429.
- Rolls, E. T. (2007). An attractor network in the hippocampus: Theory and neurophysiology. *Learning and Memory*, 14, 714–731.
- Rolls, E. T. and Xiang, J. Z. (2005). Reward-spatial view representations and learning in the primate hippocampus. *The Journal of Neuroscience*, 25, 6167–6174.
- Shapiro, M. L. and Ferbinteanu, J. (2006). Relative spike timing in pairs of hippocampal neurons distinguishes the beginning and end of journeys. *Proceedings of the National Academy of Sciences of the USA*, 103, 4287–4292.
- Shergill, S. S., Brammer, M. J., Williams, S. C., Murray, R. M. and McGuire, P. K. (2000). Mapping auditory hallucinations in schizophrenia using functional magnetic resonance imaging. *Archives of General Psychiatry*, 57, 1033–1038.
- VanRullen, R. and Koch, C. (2003). Is perception discrete or continuous? *Trends in Cognitive Sciences*, 7, 207–213.
- Vignal, J. P., Maillard, L., McGonigal, A. and Chauvel, P. (2007). The dreamy state: Hallucinations of autobiographic memory evoked by temporal lobe stimulations and seizures. *Brain*, 130, 88–99.
- Zeman, A. (2001). Consciousness. *Brain*, 124, 1263–1289.
- Zhang, Z. J. and Reynolds, G. P. (2002). A selective decrease in the relative density of parvalbumin-immunoreactive neurons in the hippocampus in schizophrenia. *Schizophrenia Research*, 55, 1–10.

The visual unconscious

Perspectives from the Charles Bonnet Syndrome

Dominic H. ffytche

Neurobiological accounts of consciousness typically partition brain activity into two components: one correlated with conscious experience (the neural correlate of consciousness) and the other not. Here I use evidence derived from visual hallucinations in the context of eye disease – the Charles Bonnet Syndrome – to examine the nature of what might be termed the visual unconscious. Forcing us to reconsider the nature of the unconscious, this hidden system of modular processing underlies our apparently seamless conscious visual experience of the world with many of its complex functions yet to be recognised by visual science.

Introduction

We open our eyes and see the world around us, a conscious visual experience of colours, objects and movement and visual spatial relations. In this context *conscious* refers to the subjective visual experience of ‘seeing’ and not the level of arousal (consciousness as the spectrum of awake, drowsy and sleep states) or awareness of self or identity (self-reflective consciousness) and is distinct from *reporting* what is seen (access consciousness, see Block, 2005). How neural inputs from the eye result in such ‘seen’ conscious experience is a question that poses a major challenge to philosophy and neuroscience alike, going beyond vision to the heart of how brains and minds interact. However, in the last decade a number of small advances have been made by focussing research on brain activities that *correlate* with conscious visual experience – the so called neural correlate of consciousness (NCC). Such studies rest on the assumption that, if consciousness resides in the brain, some brain activities must be related to conscious experience and may be revealed using imaging techniques and careful study design. Several groups have identified candidate activities and processes. Yet, while the neuroscience of visual consciousness has moved forward there has been little advance in our understanding of its natural counterpart: visual activity that might be termed

the neural correlate of unconsciousness (NCUC). Like consciousness and conscious, the terms unconsciousness and unconscious require further qualification. In the context of the NCUC unconsciousness is not used to refer to a specific state of non-arousal but to activities that are 'not conscious' or, more specifically, activities that even when fully wake and with the full allocation of visual attention are not associated with conscious visual experience.

The unconscious is familiar to all through the early 20th Century contribution of the psychodynamic tradition. The self and conscious experience are thought to rest on a hidden unconscious world of drives, wishes and conflicts with pathology of the self and consciousness reflecting changes in this inaccessible inner space. The question arises whether unconsciousness in this psychodynamic sense bears any resemblance to unconsciousness in the sense of contemporary work on the NCC and its non-conscious NCUC counterpart. Here I seek to examine this issue in the context of visual hallucinations – conscious visual experiences that occur *de novo*. I argue that in the context of eye disease, visual hallucinations provide us with a glimpse of the visual unconscious; not an unconscious as conceived by the psychodynamic tradition but instead an unconscious related to the hidden workings of our visual brains.

Visual consciousness

Based on the assumption that it makes neurobiological sense to study consciousness in different domains (visual consciousness, auditory consciousness, tactile consciousness etc.), imaging studies over the last decade have provided clues as to the nature of visual consciousness. Such studies attempt to dissociate conscious visual experience from visual input and examine differences in brain activity occurring with and without visual consciousness. As illustrated in Figure 1a, the typical finding in such studies is an increase in the level of activity within specialised visual areas (solid line) when subjects have conscious visual experiences compared to when they have not (dashed line). For example, Moutoussis and Zeki (2002) have shown that activity within cortex specialised for faces (Fusiform Face Area) increases when subjects have a conscious visual experience of a face and that activity within cortex specialized for houses (Parahippocampal Place Area) increases when subjects have a conscious visual experience of a house, while Pins and ffytche (2003) have shown that activity within early visual areas (V1 or V2) increases when subjects have a conscious visual experience of a simple visual stimulus (a contrast grating). This increment in activity constitutes a neural correlate of visual consciousness that, in some way, links to the phenomenological experience of seeing the visual attribute that the cortical region is specialised for.

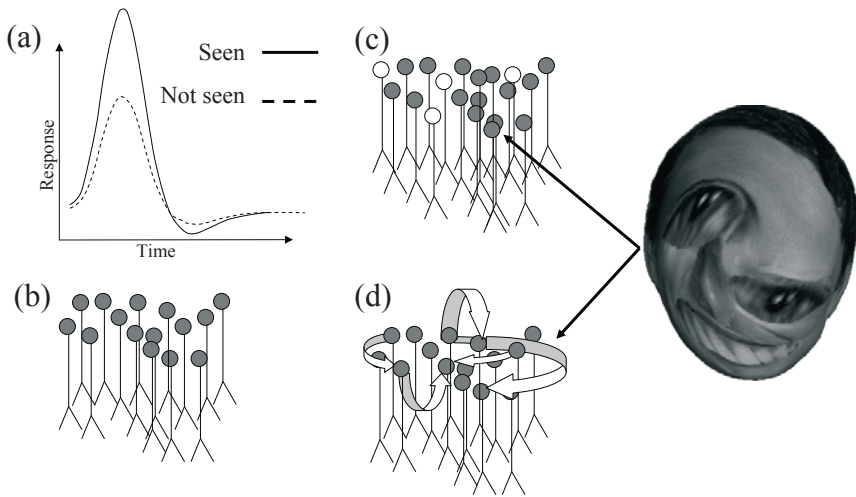


Figure 1. The NCC and NCUC in specialised visual areas (see ffytche, 2002). (a) The typical imaging finding in studies investigating neural correlates of consciousness. An increased BOLD signal (solid line) is found within a specialised visual area when subjects perceive the stimulus under investigation compared to when they do not (dashed line) under conditions in which visual inputs remain constant. (b) The NCUC, a population of cells whose activity corresponds to the dashed line BOLD response in (a) and is not associated with conscious visual experience. (c) One account of the NCC. An additional set of cells (white) become active when subjects consciously perceive the stimulus accounting for the increased BOLD signal in (a) (solid line). (d) An alternative account of the NCC. A different type of neural processing (white arrows) becomes active when subjects consciously perceive the stimulus accounting for the increased BOLD signal in (a) (solid line). The white activity illustrated in (c) or (d) underlies hallucinations of the visual experience that the area is specialised for. In the example given, the specialisation of the area is for face features and the associated hallucination is of a face dominated by features (see text).

What the increase in activity in such studies reveals of the underlying neurophysiology of conscious vision is unclear. It might reflect the recruitment of a particular neuronal cell type within a specialized area (Figure 1c white cells) or a different type of processing within the same population (Figure 1d white arrows) (see ffytche 2002). Figures 1b–d highlight the fact that a population of cells does not correlate with consciousness. Indeed, much of the early work on visual specialization was carried out in anaesthetized non-human primates where the anaesthetic ruled out a direct link to consciousness. This activity is the NCUC and we know relatively little about it other than that it carries the same specialisations as the

NCC with which it is co-localised. Whether the NCC within a given specialised area is in itself sufficient for the conscious visual experience of that specialisation remains debated, some authors arguing that downstream activity within a global workspace of frontal, parietal and cingulate areas or re-entrant inputs to the primary visual cortex are necessary (see for example Dehaene and Changeux, 2004). However, there are grounds for believing that such down stream activity is not necessary (see below).

Charles Bonnet and his Syndrome

Sometimes we see things not present in the world around us. This type of experience – a visual hallucination – is common in a number of clinical conditions affecting the brain or eye and, in the normal population, at the margins of sleep. The association of visual hallucinations and eye or visual pathway disease is today referred to as the Charles Bonnet syndrome (CBS – see ffytche, 2007 for a history of the term and its other uses). A botanist and entomologist of international repute, Charles Bonnet (1720–1793) changed career in his 30's when no longer able to use a microscope due to visual disturbances. His research turned to what would today be called the cognitive sciences and in 1760 he developed a neuroanatomical account of the conscious mind.¹ His account was based on emerging evidence of the neuronal organisation of the brain and a hypothetical neurophysiology of neural vibrations that underlay mental experience. Bonnet envisaged that the vibration of specific nerve fibres led to specific sensory experiences as well as memory, dreams, and imagery of these experiences, and that chains of mental associations were formed by the transmission of vibrations from one nerve fibre to another. To illustrate his theory he gave as an example the hallucinations² of his elderly, visually impaired grandfather. He argued that when activated inappropriately, nerve fibres with the neurophysiological imprint of a specific visual experience gave rise to spontaneous visual percepts of that experience.

CBS typically begins in the days and weeks following a significant decrease in vision. The hallucinations occur as a series of discrete episodes lasting seconds or minutes with, initially, many such episodes every day but a gradual reduction in their frequency over time. CBS hallucinations are silent, appear externally in the world and are not under volitional control in that the patient is unable to make them come or go by force of will. Patients with CBS feel as if they are spectators watching visual events unfold before them. Depending on the degree and extent of their visual loss, they will not be able to see real objects in the world. In contrast, their hallucinations are seen in perfect detail. CBS patients invariably learn

that such experiences are hallucinations; yet the hallucinations may be so compelling that the patient is often uncertain of whether a given visual experience is real or not. CBS hallucinations range from simple colours, lines and shapes, through patterns, grids, checkerboards and tapestries to complex forms. In our survey of CBS the commonest complex hallucinations were faces (typically distorted or gargoyle-like with prominent eyes and teeth), figures (typically small, wearing costumes and hats), branching forms (like the roots of a tree), letter strings (which subjects are unable to read as words), and vehicles (see Santhouse et al., 2000). Subjects also saw multiple copies of the same object (polyopia) and objects that moved with their eyes (visual perseveration).

There are no specific psychodynamic theories of CBS hallucinations. However, the traditional account of CBS as the *release* of stored memories has strong links to psychodynamic theory. Release theory was first developed by West in the 1950s (see West, 1975) who held that a constant stream of unconscious background information-processing³ was kept outside the sphere of consciousness by a process conceived in terms of scanning, filtering, and inhibiting, and linked to sensory input. Loss of sensory input released redundant information which, given a sufficient level of cortical arousal, appears in consciousness as a hallucination, the higher the arousal the more vivid the hallucinatory experience. West illustrated the theory by analogy to a man looking out of a window from a room containing a fire. In bright sunlight the man sees only the world outside; however, as night begins to fall the man begins to see things inside the room reflected on the glass. When dark outside, the man sees the contents of the room as if they were outside the window while the fire burns brightly, yet he sees nothing more as the fire dies down. In West's analogy, daylight is akin to visual sensory input, the contents of the room are akin to unconscious memories and firelight is akin to cortical arousal.

CBS hallucinations and the NCUC

The conscious experience of seeing something that is not there is exactly the same as the day-to-day experience of seeing the world around us, the only difference being that the experience occurs without an appropriate visual input. Visual hallucinations thus provide a way to study the neural mechanism of visual consciousness without the confound of visual inputs (see ffytche, 2000; ffytche et al., 1998). Our studies of CBS found that during hallucinations of colours, activity increased in colour-specialised cortex, during hallucinations of faces, activity increased in face-specialised cortex and so forth. A more recent study of visual hallucinations

has replicated the finding in the context of schizophrenia (see Oertel et al., 2007). The fact that frontal areas do not seem to be activated in such studies provides evidence that frontal cortex is not a necessary prerequisite for conscious vision, arguing against global workspace accounts of consciousness (see ffytche et al., 1998).

How do increases in activity within specialised cortical areas during hallucinations relate to the NCC and NCUC? In terms of Figure 1, it is activity in the NCC (white cells or process) that underlies the hallucination. What is unclear is where the NCC activation originates. One possibility is that it arises first within the NCUC with the transition from NCUC to NCC occurring as normal. In this case the core underlying pathology of a hallucination is spontaneous activation of the NCUC. An alternative possibility is that activity arises *de novo* within the NCC itself. Here the core pathology is of the cells or processes of the NCC, without the necessary involvement of the NCUC. Whether hallucinations reflect core pathology within the NCC, NCUC or both, they constitute a perceptual record of activity within a given specialised visual area and it follows that for every type of hallucination there must be an underlying NCC / NCUC.

For some types of visual hallucination this comes as little surprise. We know there are regions of the brain specialised for colour accounting for why patients with CBS report hallucinations of colours. Similarly we know there are regions of the brain specialised for objects accounting for why patients with CBS hallucinate objects. But what of hallucinations of distorted faces, vehicles or figures in costume wearing hats. Could cortical specialisations account for such experiences? Where this seemingly unlikely possibility has been tested, it turns out to be correct. Figure 2a shows the results of an unpublished study we have undertaken to examine the responses of normal-sighted subjects to faces distorted in such a way as to match the hallucinations of patients with CBS. The distorted face stimulus activates an area of cortex on the lateral occipital surface (the occipital face area or OFA) thought to be involved in the processing of face features. Surprisingly it activates this area better than a normal face. The larger activation evoked by the distorted face than the normal face does not seem to reflect a general difficulty of processing distorted stimuli as an object distorted in a similar fashion did not cause a similar increase in activity in this, or any other, brain region when compared to the same object without distortion. Spontaneous activation of the NCC / NCUC in the OFA would be expected to result in the hallucination of a face dominated by its features, accounting for the prominent eyes, teeth and feature distortions in the gargoyle-like face hallucinations of CBS. In support of this view, a patient with a lesion encroaching on the same lateral occipital region (Figure 2b) perceived distortions of real faces during seizures affecting this region, an experience with obvious phenomenological links to the distorted face hallucinations of CBS. One might envisage that the same principle applies for hallucinations of

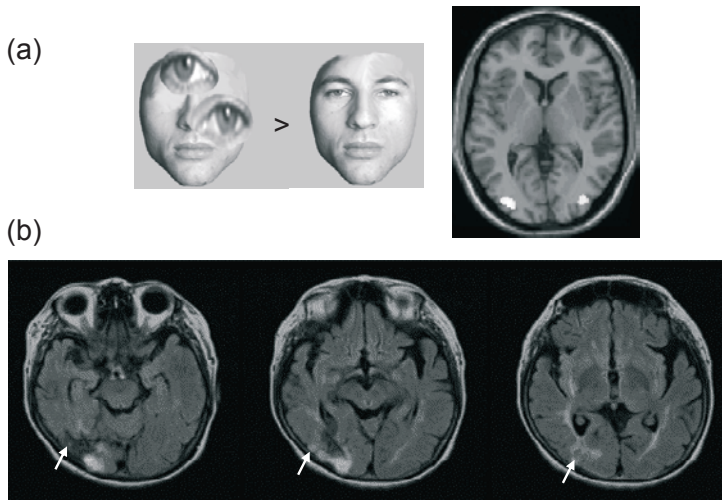


Figure 2. Brain activity related to distorted faces (unpublished data). (a) 8 normal-sighted subjects were presented a face with varying degrees of distortion and a canonical or distorted object (a teapot). The normal and maximally distorted face stimuli are shown on the left. The distorted face led to greater activation of bilateral regions in the lateral occipital lobe compared to the normal face, shown here as white areas superimposed on a high resolution axial MRI slice of brain (white regions are significant at $p < 0.05$ corrected for multiple comparisons). (b) The same lateral occipital region is active in the right hemisphere of a patient who saw real faces as distorted. The activation (white arrows) lies on the margins of a lesion in the right occipital lobe (the left of each axial brain slice is the right of the brain) (see Heo et al., 2004).

figures in period costume, vehicles, multiple copies of an object and other types of bizarre visual experience. As yet unknown specialisations of object, figure or spatial processing mean that when their corresponding NCCs become active, hallucinations occur of these seemingly bizarre categories.

Refining the visual unconscious

The observations and evidence outlined above allow us to reconsider what is meant by the visual unconscious. CBS visual hallucinations seem better conceived in terms of the generic representations, visual templates or visual processes embodied within specialised visual areas than re-activated unconscious visual memories. Indeed, there is little evidence to support a direct link between hallucination content and reactivated memory. In one study, 77% (46/60) of CBS patients

reported no personal relevance of their hallucinations with, of particular interest here, one patient even noting that their hallucinations carried less personal relevance than their dreams (see Teunisse et al., 1996). Indeed, one might argue that much of the bizarre content reported by patients with CBS could not have been encountered previously or constitute a memory in the senses of a previously encountered visual experience.

Traditional psychodynamic theory treats hallucinations as a single type of symptom irrespective of whether the hallucination is heard or seen. Freud argued that hallucinations reflect the translation of memories and ideas into formal percepts (the same process as occurs in dreams and infancy) as the result of regression to an earlier state (see Freeman et al., 1965). For example, a patient regretting the loss of contact with her family has visual hallucinations of photographs of her relatives, the conscious visual manifestation of an unconscious distressing idea or conflict (see Freeman et al., 1965). Although such accounts may help explain hallucinations in some clinical contexts they do not seem to account for CBS. What type of conflict might explain seeing a distorted face as opposed to a tapestry pattern or a formless colour? Such experiences seem better accounted for in terms of the mechanisms of visual processing. Indeed, one psychodynamic theory holds that bizarre visual hallucinations are not specific ideas or conflicts but to relate instead to archetypes⁴ (see Faergeman, 1963), an account which could be conceived as a metaphor for collectively shared mechanisms of vision embodied in our brains.

When viewed from the perspective of CBS the visual unconscious can be conceived as a hidden realm, not of conflicts, drives wishes or memories as held by psychodynamic theory, but of visual processing. This realm is hidden in two distinct ways. First NCUC activity does not have a conscious correlate even in states of full attention and arousal and is thus hidden by definition. CBS hallucinations provide indirect views of the NCUC through its co-localised NCC; yet, hallucinations are not unique in this respect. Our current understanding is that all day-to-day visual experience is based on co-localised NCUC and NCC activity so that all conscious vision provides a view of the NCUC. The key difference between CBS hallucinations and day-to-day vision is that in CBS the NCC and NCUC activities arise out of context and thus reveal a second hidden aspect of visual processing – its modularity. Day-to-day visual experience gives no clue as to the modular nature of visual processing and we are unable to overcome by force of will the impression of a seamless, unitary visual experience of colours, motion form and spatial location. This modularity is not amenable to introspection and can be considered part of the visual unconscious although unconscious in this context need not imply the NCUC as it remains unclear whether such unitary experience has an underlying neurobiology or is illusory. By placing modules out of their normal context, CBS hallucinations betray their unsuspected presence

in our day-to-day visual experience. Take, for example, distorted face hallucinations. Even though we have a separate NCC for face features and face wholes, our subjective experience of seeing faces is not of two separable face qualities (features and wholes) but of a face, unified with other visual attributes and sense modalities. By placing the face feature module within consciousness but out of its normal context, a distorted face hallucination reveals the presence of an otherwise unsuspected specialisation. CBS hallucinations thus provide snapshots of separable visual modules, revealing a range of visual processes and functions, many of which have yet to be recognised by visual science.

Beyond Charles Bonnet Syndrome

The neurobiological account of hallucinations outlined above refers specifically to the context of eye disease. Here hallucinations are entirely visual and can thus be related to the visual cortex and its various NCC / NCUC specialisations. Can the same principles be applied to hallucinations and the unconscious in other domains? Although less is known of the auditory NCC, one might predict the same principles apply as in the visual system. Consistent with this view, both structural and functional changes are found within auditory and language circuitry in patients with auditory verbal hallucinations in the context of schizophrenia (see Allen et al., 2008). Where visual hallucinations occur in such patients, they are invariably associated with another modality, either at the same time (a multimodal hallucination) or at different times (auditory hallucinations on some occasions, visual hallucinations on others). It is therefore of interest that some of the cortical areas implicated in the visual hallucinations of schizophrenia are heteromodal with both auditory and visual modalities represented (e.g. limbic activation not found in CBS, see Oertel et al., 2007). One might hypothesise that it is the NCC in heteromodal cortex that causes the hallucinations of both sensory modalities simultaneously or the switch from one modality to another.

Conclusions

Charles Bonnet realised the significance of visual hallucination in helping understand the neural mechanisms of the mind. 250 years later the account presented here differs little from Bonnet's original conception. Replace specialised 'nerve fibres' with the 'NCC' or 'NCUC' and replace spontaneous 'vibration' with 'activation' and the two accounts are almost identical. Where we have moved forward is in a better understanding of what visual hallucinations reveal about the

architecture of the visual mind. Visual hallucinations seem to make accessible aspects of processing underlying our day-to-day visual experience that remain hidden under normal circumstances. In so doing they challenge our current models of visual processing and force us to reconsider what we mean by the visual unconscious.

Notes

1. Bonnet referred to the *Soul* but the scope and remit of his theory relates better to the modern day usage of *Mind* and *Consciousness*.
2. Bonnet used the term *Visions*. *Hallucination* was not used in its current sense until the early 19th century.
3. West referred to the information processing as 'percepts or memory traces' in his 1962 monograph (see Cogan 1973).
4. Stored ancestral memories.

References

- Allen, P., Laroi, F., McGuire, P. K. and Aleman, A. (2008). The hallucinating brain: A review of structural and functional neuroimaging studies of hallucinations. *Neuroscience and Biobehavioral Reviews*, 32, 175–91.
- Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences*, 9, 46–52.
- Cogan, D. G. (1973). Visual hallucinations as release phenomena. *Graefes' Archive for Clinical and Experimental Ophthalmology*, 188, 139–50.
- Dehaene, S. and Changeux, J. P. (2004). Neural mechanisms for access to consciousness. In M. S. Gazzaniga (Ed.). *The cognitive neurosciences* (pp. 1145–1157). Cambridge, MA: MIT Press.
- Faergeman, P. M. (1963). *Psychogenic psychoses: A description and follow-up of psychoses following psychological stress*. London: Butterworths.
- ffytche, D. H. (2000). Imaging conscious vision. In T. Metzinger (Ed.). *Neural correlates of consciousness: Empirical and conceptual questions* (pp. 221–230). Cambridge, MA: MIT Press.
- ffytche, D. H. (2002). Neural codes for conscious vision. *Trends in Cognitive Sciences*, 6, 493–495.
- ffytche, D. H. (2007). Visual hallucinatory syndromes: Past, present, and future. *Dialogues in Clinical Neuroscience*, 9, 173–89.
- ffytche, D. H., Howard, R. J., Brammer, M. J., David, A., Woodruff, P. and Williams, S. (1998). The anatomy of conscious vision: An fMRI study of visual hallucinations. *Nature Neuroscience*, 1, 738–742.

- Freeman, T., Cameron, J. L. and McGhie, A. (1965). *Studies on psychosis: Descriptive, psycho-analytic, and psychological aspects*. London: Tavistock Publications.
- Heo, K., Cho, Y. J., Lee, S. K., Park, S. A., Kim, K. S. and Lee, B. I. (2004). Single-photon emission computed tomography in a patient with ictal metamorphopsia. *Seizure*, 13, 250–3.
- Moutoussis, K. and Zeki, S. (2002). The relationship between cortical activation and perception investigated with invisible stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 9527–9532.
- Oertel, V., Rotarska-Jagiela, A., van de Ven, V. G., Haenschel, C., Maurer, K. and Linden, D. E. (2007). Visual hallucinations in schizophrenia investigated with functional magnetic resonance imaging. *Psychiatry Research: Neuroimaging*, 156, 269–73.
- Pins, D. and ffytche, D. H. (2003). The neural correlates of conscious vision. *Cerebral Cortex*, 13, 461–474.
- Santhouse, A. M., Howard, R. J. and ffytche, D. H. (2000). Visual hallucinatory syndromes and the anatomy of the visual brain. *Brain*, 123, 2055–2064.
- Teunisse, R. J., Cruysberg, J. R., Hoefnagels, W. H., Verbeek, A. I. and Zitman, F. G. (1996). Visual hallucinations in psychologically normal people: Charles Bonnet's syndrome. *Lancet*, 347, 794–797.
- West, L. J. (1975). A clinical and theoretical overview of hallucinatory phenomena. In R. K. Siegel and L. J. West (Eds.), *Hallucinations* (pp. 287–311). New York, NY: John Wiley and Sons.

Believing is hearing is believing

The reciprocal nature of consciousness

William Sedley

Auditory consciousness takes many forms. Some of these, such as auditory perception and imagination, are familiar to nearly all people, while others, such as tinnitus or hearing voices, are unfamiliar to most. This chapter proposes a simplified conceptual model of the auditory system, involving bottom-up influences from the lower auditory pathways and top-down influences from higher evaluative mechanisms. This model is used to explain auditory perception, hallucinations and illusions based, for each phenomenon, on only minor variations from the normally functioning system. Neuroimaging and electrophysiological correlates of these processes are discussed, along with their inability to qualitatively differentiate conscious from unconscious processes, and how we might approach this problem in future.

Everybody with a functioning auditory system knows what is like to hear, and thereby experiences auditory consciousness. In fact, few if any would be able to remember the last time they were awake and *not* experiencing some sort of auditory consciousness (i.e. in “perfect silence”). By contrast, auditory hallucinations (the experience of auditory consciousness in the absence of a corresponding auditory stimulus) are generally thought of as the experience of only a minority of individuals, who are often stigmatised as a result. It is therefore easy to imagine that these processes are distinctly and qualitatively different in terms of their underlying neural mechanisms. This chapter explores neural mechanisms underlying normal auditory perception, auditory hallucinations and also auditory illusions. The result is a proposed common model for auditory consciousness, which addresses issues including how a diverse range of mental processes can result in the same experience, how we can further refine our understanding of the neural correlates of consciousness and how it can be too loud to hear ourselves think. The content of this chapter is not intended to provide an exhaustive account of every neural mechanism that can lead to auditory consciousness, but rather to highlight common elements to some of the more frequently-encountered auditory phenomena.

By contrast, these processes need not apply only within the auditory system, but might be similar to those resulting in perceptual consciousness in other modalities such as the visual or olfactory systems.

Definitions

Before we begin developing arguments and models we should clarify some terms, and their assumed meanings within this chapter, that are crucial to building these. These are not intended to be unfamiliar or controversial concepts, and neither are they proposed as authoritative definitions, but simply to introduce key concepts and provide an unambiguous language for developing the arguments to be made.

“Auditory consciousness” – The subjective experience of *hearing* something. It is assumed that this entity is inextricably linked to (or entirely caused by) a certain type of neural activity within the brain, its “neural correlates”. This definition excludes all neural activity which is sometimes associated with auditory consciousness (e.g. electrical impulses being relayed from the cochlea along the auditory nerve), but that is not essential for auditory consciousness to occur.

“Auditory perception” – Auditory consciousness arising principally as a consequence of an acoustic stimulus being relayed via the auditory pathways, where the resulting conscious experience is an *accurate* representation of the causative stimulus.

“Auditory illusion” – Auditory consciousness arising principally as a consequence of an acoustic stimulus being relayed via the auditory pathways, where the resulting conscious experience is an *inaccurate* representation of the causative stimulus (i.e. one thing is *heard* but a different thing is *perceived*, such as the false perception of hearing somebody say your name in a noisy room). It is fairly obvious that no absolutely concrete division can exist between this phenomenon and normal auditory perception, but the distinction between the two concepts is helpful.

“Auditory imagery” – A weaker, qualitatively different form of auditory consciousness; one has an awareness of a sound or voice and its acoustic characteristics, but without actually *perceiving* it to its normal intensity, or as if it is heard from an external source. Examples include thought, memory, imagination and silent reading.

“Auditory representation” – A pattern of neural activity in the auditory cortex that is an encoded representation of an auditory stimulus, but that is separate from the neural correlates of consciousness that might result from it.

“Externality label” – A label, signature or property attached to an auditory representation that cause it to be experienced as full auditory consciousness (as if perceived from the environment) rather than auditory imagery or unconscious auditory activity. It does not necessarily follow that the externality label is itself a

neural correlate of consciousness, as it could instead be a pattern of neural activity that stimulates the neural correlates of auditory consciousness.

“Unconscious auditory activity” – Similar in its origins to auditory imagery. Processes utilising the auditory cortex to generate auditory representations (reading, speech planning, subconscious thought, etc), but where there is no conscious awareness of this representation either as auditory perception or auditory imagery. i.e. The person is completely unaware of this auditory representation taking place.

“Auditory hallucination” – Auditory consciousness, perceived as if heard from an external source, arising from mechanisms other than as a consequence of sound being relayed through the auditory pathways (e.g. hearing voices talking about you in schizophrenia). In this chapter, the term also covers tinnitus (the perception of “ringing”, “buzzing” or other droning noises in the ear(s)) and musical hallucinations.

“Auditory cortex” – The part of the cerebral cortex in the temporal lobes that processes, interprets, remembers, recalls and imagines sounds. For the purposes of this chapter, there is not generally drawn a distinction between primary, secondary and association auditory cortex. While it is understood that there is hierarchical processing of sounds, from simple to increasingly complex characteristics, from primary to association auditory cortex, this organisation is largely irrelevant to the arguments made herein. I also wish to avoid the vogue for trying to define the responsible anatomical “centre” for various brain functions and conscious perceptions. While this approach has many merits, it can be reasonably argued that it simply shrinks the box in terms of defining consciousness rather than actually helping to delineate its underlying mechanisms.

“Evaluative mechanisms” – Communications from the auditory cortex to parts of the brain involved in semantic evaluation, emotional responses, auditory memory, cross-modal integration and interaction with the rest of the conscious and subconscious mind. This includes the frontal, parietal and temporal cortices, subcortical structures such as the thalamus, amygdala and hippocampus, and parts of the auditory cortex itself. In essence, these are higher brain functions that are fed by auditory consciousness, imagery and unconscious auditory activity but are not directly responsible for supporting auditory consciousness or encoding other auditory representations.

“Predictive mechanisms” – Communications to the auditory cortex for the purposes of generating auditory imagery (e.g. thought, speech rehearsal, music/lip reading), recalling from auditory memory (e.g. predicting the next note in a tune or word in a sentence) or tuning the perception of externally-perceived sounds (e.g. hearing yourself sing the note you intended rather than the one you actually produced). These communications can be from many of the brain regions associated with evaluative mechanisms, or from the auditory cortex itself.

The auditory perceptual pathways

The (simplified) model of normal auditory perception proposed in this chapter is illustrated in Figure 1a. The other parts of Figure 1 apply the same model to other forms of auditory representation, consciousness and hallucinations. The following paragraphs describe and explain these model variations.

Auditory perception (Figure 1a)

Auditory stimuli are encoded into nerve impulses by the cochlea. This electrical activity undergoes processing in the early auditory pathways in the brain stem and medial geniculate body. Early processing includes gain whereby, in low sound conditions, stimuli are amplified to make them more salient and therefore more amenable to interpretation (the auditory equivalent of night-vision). Activity is relayed to the auditory cortex, which results in auditory representations. Competitive processes take place in the auditory cortex, whereby representations are evaluated in terms of salience, relevance, plausibility, consistency and expectedness based on concurrent or preceding stimuli (for a discussion of the issue of *auditory objects* and how these are processed, see Griffiths et al., 2004). Higher-scoring representations are more strongly represented, and lower-scoring ones disintegrate and disappear. Only the highest-scoring representations reach consciousness, and the threshold for this to occur depends on factors such as the number and salience of other current representations, overall wakefulness and selective attention being paid. However, it is not simply the case that the highest scoring representation is consciously perceived, as we do not hear out thoughts spoken aloud simply by sitting in a very quiet place. The described competitive processes occur via feedback loops within the auditory cortex itself, and through a wider feedback loop involving communication with evaluative mechanisms, which reciprocate with the auditory cortex by way of predictive mechanisms. It has been shown that conscious processes are strongly associated with cortical electrical oscillations of 40 to over 100 Hz (gamma oscillations), and that these display synchrony (i.e. communication) both locally and between anatomically separate brain regions involved in the processing of a given percept (for a discussion of gamma oscillations in general and with relevance to auditory perception see Tallon-Baudry et al., 1999 and Crone et al., 2001 respectively). Given that conscious perception develops over a delay of hundreds of milliseconds, it is probable that a large number of feedback cycles have occurred (and thereby extensively modified an auditory representation) before any auditory consciousness begins to occur. This would explain why we usually perceive auditory percepts as fully formed from the outset (e.g. hearing a doorbell ring or a word spoken rather than a collection of tones and sounds that subsequently assemble themselves into a singularly perceived entity).

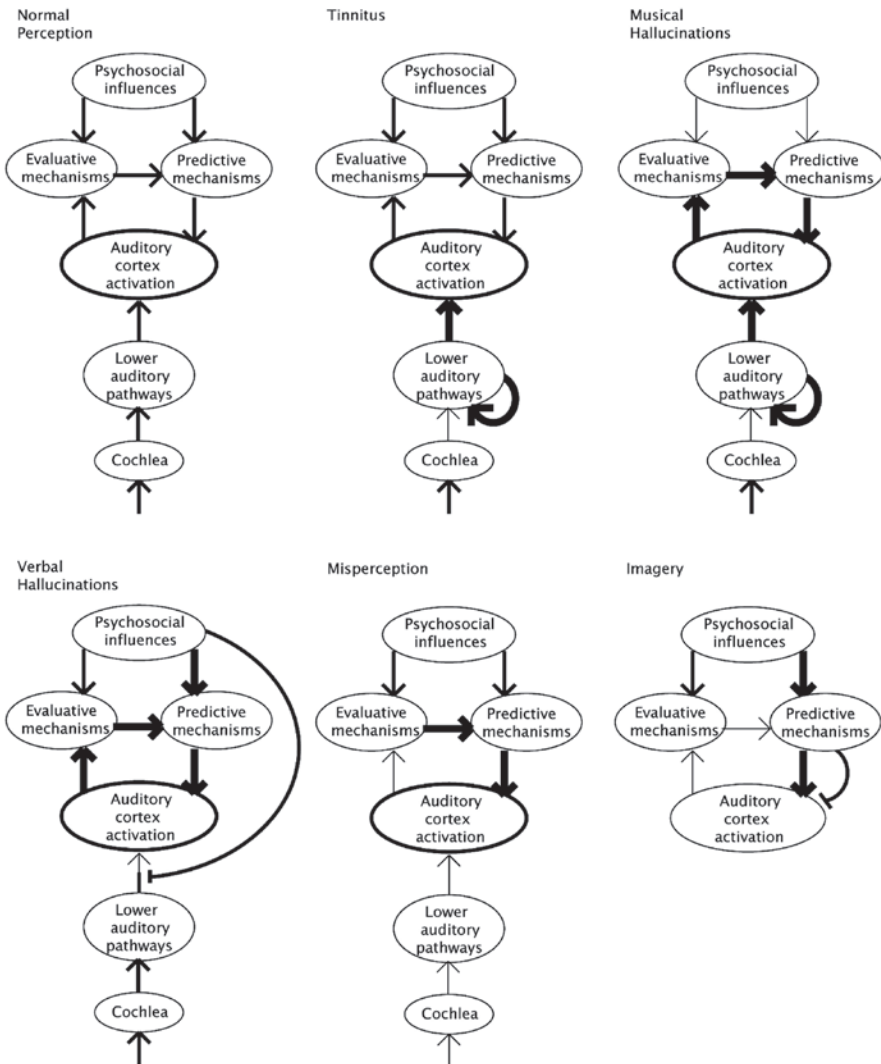


Figure 1. Simplified model of the auditory perceptual pathways, and how these are modified to produce all types of auditory consciousness. Standard thickness arrows denote connections of normal strength, with thinner or wider arrows denoting reduced or increased connectivity respectively. Blunt-ended arcs denote inhibitory connections.

Tinnitus (Figure 1b)

Stimuli reaching the cochlea are unchanged from normal. However the cochlea has suffered damage, resulting in high-frequency hearing loss. Therefore regions of the lower auditory pathways that receive input from these damaged parts of the cochlea are understimulated. This understimulation leads to spontaneous activity, which is subject to gain mechanisms, that causes activation of the auditory cortex and thereby auditory representation and consciousness (see Eggermont et al., 2004 and Llinás et al., 2005 for discussion of how understimulation leads to spontaneous activity, Lockwood et al., 1998 for demonstration of human auditory cortex activation and Ashton et al., 2007 for gamma oscillations in tinnitus). Feedback through evaluative and predictive mechanisms is not an essential part of the mechanism of tinnitus generation. However negative emotional reactions to the perception of tinnitus are believed to cause hypersensitivity of the auditory cortex to the tinnitus percept, and therefore, via a positive feedback cycle, a more intense experience of tinnitus and the tendency to selectively focus on the tinnitus above externally perceived sounds (see Jastreboff, 1990 for a model of tinnitus that continues to guide successful psychological intervention today). However, hearing loss is not always the initiating event in the generation of tinnitus. Ingestion of salicylates, withdrawal from benzodiazepines and sensorimotor stimuli from the face, head and neck (Levine et al., 2003) can all lead to or enhance tinnitus. The likelihood is that tinnitus represents an inherent tendency for the auditory system to produce, enhance and interpret activity, which becomes pathologically overactive under certain conditions.

Musical hallucinations (Figure 1c)

Similarly to in tinnitus, there is significant hearing loss at the level of the cochlea, and therefore spontaneous activity in the lower auditory pathways. In contrast to tinnitus, in this condition there is another contributing factor, possibly social isolation, which influences how the spontaneous auditory activity is interpreted. For instance, social isolation might leave cognitive and emotional evaluative networks understimulated through normal channels, leaving them hyper-reactive to stimulation by the auditory cortex. When these evaluative mechanisms are engaged, they feed back to the auditory cortex, via predictive mechanisms, and modify the auditory representation accordingly to make it fit the preferred interpretation (i.e. music). Once the necessary conditions are fulfilled for the spontaneous auditory activity to begin to be perceived as music, its interpretation as music by evaluative mechanisms becomes strengthened, and these increase their feedback to predictive mechanisms and auditory cortex. A positive feedback cycle thus quickly develops. Due to the strong sequential nature of the perceived music, once a short section of the melody (even, perhaps, one or two notes) becomes established, the

rest easily follows on, as each section is predicted from the last by the very cognitive mechanisms that initiated the process (see Griffiths et al., 2000 for a comprehensive model for musical hallucinations). While hearing loss and resultant understimulation is the principal cause for musical hallucinations, it should be bourn in mind that other processes (such as focal seizures and psychotic illness) can lead to musical hallucinations.

Verbal hallucinations (Figure 1d)

The model discussed below is primarily based around hallucinations in schizophrenia (which most classically take the form of hearing one's thoughts spoken aloud or hearing voices talking to or discussing one's self), but it is intended to not be incompatible with verbal hallucinations experienced by other mechanisms. Unlike the above examples of tinnitus and musical hallucinations, the primary process underlying verbal hallucinations is a top-down one. Predictive mechanisms that usually result in auditory representations or imagery (see Shergill et al., 2000 and 2003 for the relationship between verbal imagery and verbal hallucinations), but not perception, (e.g. verbal thoughts or silent articulation) become detached from corollary discharges that normally suppress resultant auditory consciousness by preventing the generation of an externality label. i.e. In the normal situation auditory imagery is co-generated and co-transmitted with activity that identifies it as internally-generated, and that this process is disrupted in the brains of people with schizophrenia who are prone to verbal hallucinations. Experimentally, normal subjects have been shown to suppress auditory cortex reactivity when engaging in verbal imagery, but people with schizophrenia who are prone to verbal hallucinations do so to a much lesser extent (see Ford et al., 2005). Consistent with this finding, schizophrenia has been found to be associated with disruption of synchronised gamma oscillations between brain regions (see Spencer et al., 2003), which might explain the dysfunction of corollary discharges. An additional possible influence is that the stressful life circumstances that often precipitate the onset of psychosis result in a partial detachment from environmental influences (in this case auditory perception), which therefore leave the auditory cortex understimulated, and prone to internally-generated stimulation (through reduced competition) as described above. Once the threshold is crossed where auditory consciousness can result from internally-generated processes, a positive feedback cycle is created (via evaluative and predictive mechanisms) where the perceived speech is concordant with the current content of thought (because the latter produced the former), and therefore each strengthens the other.

Misperception (Figure 1e)

A limited amount of auditory information is conveyed to the auditory cortex, either because the stimulus is quiet, or it is perceived alongside a high level of background noise. Sometimes there might not actually be a “stimulus” as such, but rather only a minority of the afferent signal might be subjected to higher interpretation. Whichever the case is, the auditory representation in question is of insufficient salience, quantity and/or clarity to permit easy recognition. This results in increased interpretation by evaluative and predictive mechanisms which, when correctly applied, would result in optimum stimulus interpretation. However in these situations of low stimulus information, such mechanisms rely more heavily on prediction based on other thought or stimuli. This provides a mechanism whereby the content of thought can be projected into auditory consciousness provided there is a rough concordance between the content of thought and the afferent stimulus. In most situations where the content of thought is projected into the auditory domain (e.g. auditory imagery), it does not gain an externality label, and is hence not consciously perceived. However in the case of misperception, the content of thought is projected onto auditory cortical activity that already has an externality label, resulting in full auditory consciousness rather than imagery.

Auditory imagery and unconscious auditory representation (Figure 1f)

These phenomena essentially arise from the same process, with the difference that unconscious auditory representations result from weaker input from predictive mechanisms or fare less well in competitive selection in the auditory cortex. In either case, predictive mechanisms feed into the auditory cortex and generate an auditory representation. This is accompanied by corollary discharges that prevent an externality label from being gained. The result is that although there is feedback via evaluative and predictive mechanisms, this is by definition not sufficient to sustain or amplify the auditory representation. This means that the auditory representation cannot become strong enough to reach full consciousness, and it will only endure as long as there is continuing external input to predictive mechanisms that are driving it. This is essential from an evolutionary point of view, as an organism that became capable of satisfying its perceptual wants by imagination alone would lose the drive to continually engage with its surroundings that is necessary for survival. One of the most enduring examples of auditory imagery is having a “tune stuck in your head”, and even then it never reaches full consciousness, and cannot begin to compete with more salient or externally perceived auditory representations. Input from lower auditory pathways has been omitted from the diagrams to indicate that its presence or absence is not necessary for this phenomenon. However, as alluded to, these processes occur most easily in the relative

absence of external auditory stimuli, or by selectively ignoring external stimuli. The later situation will be familiar to most people, who will at times found themselves “unable to think” due to excessive background noise.

Prerequisites for auditory consciousness

The above models all describe the generation of auditory representations, and in most cases these lead to auditory consciousness (either weak as in auditory imagery, or full consciousness as in perception, misperception and hallucination). Neuroimaging studies on all of these processes demonstrate associated activity in the auditory cortex (see individual references listed under each process), and there is no clear, consistent difference in the anatomical distribution of this activity that would suggest one or more given locations or “centres” which are active in auditory consciousness, less active in auditory imagery and inactive in unconscious auditory representations (i.e. “auditory consciousness centres”). Furthermore, it would seem redundant and inefficient for the brain to develop parallel representational networks of the kind that would allow this (i.e. one neuronal assembly to perceive a stimulus, and another, separate, equivalent assembly to recall or imagine it subsequently). Rather, I argue that auditory consciousness is an emergent property of the auditory system, which arises when the following conditions, or prerequisites, are satisfied:

Auditory representation

The auditory cortex needs to be stimulated from the lower auditory pathways (as in normal perception or tinnitus) and/or predictive mechanisms (as in imagery or verbal/musical hallucinations) in order to activate the neuronal assembly that encodes an auditory representation.

Externality label

The externality label (or lack of an “internality label”) is always present when an auditory representation arises as a result of input from the lower auditory pathways, but is necessarily absent from auditory representations arising from predictive mechanisms. Without it, only weak auditory consciousness (i.e. imagery) can occur. In verbal hallucinations, the externality label is falsely attributed due to failure of corollary discharge mechanisms that would usually prevent this occurrence.

Stability/reciprocity

All of the above examples of auditory consciousness have a feedback cycle inherent to their underlying mechanism. In tinnitus, this is the self-sustaining activity

in the lower auditory pathways, resulting in persistent activation of the auditory cortex. In other forms of auditory consciousness (and often also in tinnitus) it arises from a reciprocal loop consisting of auditory cortex, evaluative and predictive mechanisms. It is this loop that in normal auditory perception fine-tunes perceived sounds into fully-formed, coherent percepts prior to entering auditory consciousness. In this situation the loop cannot be sustained once the input from the lower auditory pathways ends. The same loop permits imagery in the absence of an auditory stimulus, but cannot self-sustain without continual input to predictive mechanisms. In the presence of an inappropriately attributed externality label, these loops can self-sustain and self-amplify, resulting in auditory hallucinations.

Competitive weighting

Even with the first three conditions satisfied, an auditory representation will only make it into consciousness if it is selected, through competitive processes via reciprocal loops as previously described, above other auditory representations. A very loud or sudden stimulus might quickly surpass the threshold necessary to enter auditory consciousness, while a quiet stimulus against background noise might only do so after a delay (if at all) while its competitive weighting is enhanced through evaluatory mechanisms. In the case of verbal hallucinations, the competitive weighting of the initially generated (subconscious) auditory representation is extremely low, and nowhere near the threshold for entering consciousness. As reciprocal feedback cycles occur, auditory representations that are going to become hallucinations must go through a very high number of cycles, and their competitive weighting often increases in an exponential fashion (in contrast to the majority of internally-generated representations that will not endure beyond a few cycles). By the time the threshold for conscious perception is reached, the auditory representation is often so stable and has such a rapidly-increasing competitive weighting that there might be no time for voices to start as whispers, or to fade in gently as intermittently perceived words or phonemes. Instead they can arrive loud and clear, and as salient as any externally-perceived voice. In other situations, a more tenuous balance of processes enhancing or suppressing verbal hallucinations might exist, allowing for a gradual fading in of hallucinatory experiences.

Implications for the study of consciousness

The emergent nature of consciousness

We can see from the above models that similar activity within the same regions of auditory cortex can give rise to anything from no conscious awareness to a strong conscious awareness. It is therefore unlikely that functional imaging studies looking at which brain regions “light up” in association with various cognitive or perceptual tasks will shed much more light on consciousness. However, there is potentially a lot to be gained by the qualitative analysis of the type of activity that occurs within a single brain region for an equivalent auditory representation in its unconsciously represented, imagined, perceived and hallucinated forms. For instance, reading a phrase, silently articulating it, imagining a voice speaking it or being played a recording of a voice speaking it.

The reciprocal nature of consciousness

Feedback loops between auditory cortex, evaluative and predictive mechanisms are a necessary part of auditory consciousness, but also of auditory imagery and unconscious auditory representation. The findings of gamma-band oscillations as a correlate of consciousness are therefore likely to signify their importance in facilitating reciprocal connectivity between cortical regions, but does not imply that gamma oscillations and the neural correlates of consciousness are one and the same. For instance, gamma band oscillations have also been reported as a correlate of increased selective attention without a change in conscious perception (see Ray et al., 2008). If my assertion that multiple executions of these cycles often need to occur before consciousness arises, it should in theory be possible to separate gamma oscillations subserving pre-conscious representation from those subserving conscious perception in order to identify neural correlates unique to consciousness itself.

Summary

In this chapter we have introduced a simplified model of the auditory perceptual pathway, and shown how this single network (even in its simplified form) can give rise to the whole spectrum of normal and abnormal auditory consciousness. The process underlying these forms of consciousness, and unconscious auditory representation, seems to be reciprocal connectivity with evaluative mechanisms, which can be measured as synchronised gamma band oscillations in neuronal electromagnetic activity. All types of auditory consciousness seem to result from activity within the same anatomical regions of the brain, so the key to the degree of consciousness perceived seems to be the exact nature of these reciprocal

connections and the resultant neuronal activity patterns. We have explored how all individuals can be prone to auditory illusions, and how disruptions in connectivity in the normal auditory pathways can lead to positive feedback cycles that cause and enhance hallucinations.

References

- Ashton, H., Reid, K., Marsh, R., Johnson, I., Alter, K. and Griffiths, T. (2007). High frequency localised “hot spots” in temporal lobes of patients with intractable tinnitus: A quantitative electroencephalographic (QEEG) study. *Neuroscience Letters*, 1, 23–28.
- Crone, N. E., Boatman, D., Gordon, B. and Hao, L. (2001). Induced electrocorticographic gamma activity during auditory perception. *Clinical Neurophysiology*, 112, 565–582.
- Eggermont, J. J. and Roberts, L. E. (2004). The neuroscience of tinnitus. *Trends in Neurosciences*, 27, 676–682.
- Ford, J. M. and Mathalon, D. H. (2005). Corollary discharge dysfunction in schizophrenia: Can it explain auditory hallucinations? *International Journal of Psychophysiology*, 58, 179–189.
- Griffiths, T. D. (2000). Musical hallucinosis in acquired deafness: Phenomenology and brain substrate. *Brain*, 123, 2065–2076.
- Griffiths, T. D. and Warren, J. D. (2004). What is an auditory object? *Nature Reviews Neuroscience*, 5, 887–892.
- Jastreboff, P. J. (1990). Phantom auditory perception (tinnitus): Mechanisms of generation and perception. *Neuroscience Research*, 8, 221–254.
- Levine, R. A., Abel, M. and Cheng, H. (2003). CNS somatosensory-auditory interactions elicit or modulate tinnitus. *Experimental Brain Research*, 153, 643–648.
- Llinás, R., Urbano, F. J., Leznik, E., Ramirez, R. R. and van Marle, H. J. (2005). Rhythmic and dysrhythmic thalamocortical dynamics: GABA systems and the edge effect. *Trends in Neurosciences*, 28, 325–333.
- Lockwood, A. H., Salvi, R. J., Coad, M. L., Towsley, M. L., Wack, D. S. and Murphy, W. W. (1998). The functional neuroanatomy of tinnitus: Evidence for limbic system links and neural plasticity. *Neurology*, 50, 114–120.
- Ray, S., Niebur, E., Hsiao, S. S., Sinai, A. and Crone, N. E. (2008). High-frequency gamma activity (80–150 Hz) is increased in human cortex during selective attention. *Clinical Neurophysiology*, 119, 116–133.
- Shergill, S. S., Brammer, M. J., Fukuda, R., Williams, S. C., Murray, R. M. and McGuire, P. K. (2003). Engagement of brain areas implicated in processing inner speech in people with auditory hallucinations. *British Journal of Psychiatry*, 182, 525–531.
- Shergill, S. S., Brammer, M. J., Williams, S. C., Murray, R. M. and McGuire, P. K. (2000). Mapping auditory hallucinations in schizophrenia using functional magnetic resonance imaging. *Archives of General Psychiatry*, 57, 1033–1038.
- Spencer, K. M., Nestor, P. G., Niznikiewicz, M. A., Salisbury, D. F., Shenton, M. E. and McCarley, R. W. (2003). Abnormal neural synchrony in schizophrenia. *Journal of Neuroscience*, 23, 7407–7411.
- Tallon-Baudry, C. and Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Sciences*, 3, 151–162.

Dreaming as a physiological psychosis

Connecting states of consciousness

Armando D'Agostino, Ivan Limosani and Silvio Scarone

Periodicity, hallucinatory phenomena and lack of awareness of one's detachment from reality seem to connect the subjective experience of dreaming to that of psychosis. Both can be considered fully conscious states of the brain/mind, with an immediate influence of non-conscious elements caused by an impaired interaction with the external world. Psychological and pharmacological induction of lucidity in dreams may prove useful in the comprehension of acute psychoses and in clinical practice.

In 1896 William James stated: "Sleep would be a dreadful disease but for its familiarity ... We do not regard dreaming as morbid because it is customary, but if it were not, it would be the subject of much medical wonder." (Taylor, 1983). Thus, almost 120 years ago, sleep and dreams were recognized as highly peculiar phenomena, enigmatic and somewhat pathological. The aim of this chapter is to demonstrate the plausibility of the hypothesis that dreaming mimics being psychotic in its fundamental intrinsic aspects. The validation of such hypothesis would lead to dreaming and being psychotic to be considered as the direct expression of specific functional states of the mind, thus calling for a careful examination of the problem of boundaries between the physiology and pathology of the mind. The difference between being psychotic and being affected by a psychotic mental disorder will also be carefully considered, in order to support the conclusion – that currently does not appear to be so clearly formulated in scientific literature – that the two conditions are not the same. In conclusion, the relationship between conscious and non-conscious processes will be discussed with reference to recent research on lucid dreaming with possible implications for the field of psychiatry.

The dreaming brain/mind

Dreaming has been defined as a mental activity characterized by vivid sensorimotor imagery that is experienced as waking reality despite such distinctive cognitive features as impossibility or improbability of time, place, characters and related actions (Hobson et al., 2000). The first intrinsic property of dreaming to be considered useful to our discussion is its periodicity, its recurrence in function of time. Our daily life cycle may be considered over 24 hours in terms of alternation between waking and sleeping and, within sleep, between periods in which one has the subjective experience of dreaming and others in which one doesn't. Quantitatively, this experience has a great interindividual variability: whereas some subjects report they never dream, others that they dream regularly every night. It seems reasonable to ascribe this variability to memorization mechanisms. Indeed, the experience of dreaming recalled upon awakening is commonly labile in its particulars and, in the majority of cases, tends to rapidly fade. Although dreaming appears to occur across all stages of sleep, it is broadly accepted that a subject is more likely to recall a dream when awakened at the end of a REM stage.

Several experimental settings have associated highly bizarre, vivid, hallucinoid dreams with REM sleep periods, and REM-like neurobiological modifications have been hypothesized to underlie similar subjective experiences during different EEG-defined stages of sleep (Nielsen, 2000). Dreaming usually occurs after a period of dreamless sleep, with dreams only rarely appearing at the beginning of sleep in physiological conditions. This evidence supports the hypothesis of a more than casual relationship between REM sleep and dreaming, because the same mechanism is observed in the structure of sleep, with REM stages rarely occurring at the beginning of sleep in physiological conditions.

Several theories and models have been proposed to explain this rigid cyclicity, on both the psychological and neurophysiological level. Early psychoanalysis considered the experience of dreaming as the expression of a satisfaction of unconscious desires of which dreamless sleep represents a preliminary appetitive phase (Freud, 1895). Classical neurophysiology proposed to consider sleep as an instinct, made up of a preparatory stage and a consummatory one, echoing the psychoanalytic model in some way. Considering the wakefulness-coma continuum determined by the Ascending Reticular Activating System (ARAS), the model hypothesized a first preparatory stage-consummatory stage cycle between wakefulness and sleep, where the ARAS activated/deactivated in order to induce or not to induce one of the states, and a second cycle between NREM sleep and REM sleep, within the first, in which the ARAS further deactivated (deep sleep NREM stage) or reactivated (REM stage) on the border with sleep and coma (Moruzzi, 1969). Recent research confirmed the role of the ARAS in

the genesis and maintenance of wakefulness and REM sleep, with the mutual interplay between cholinergic, serotonergic and noradrenergic systems sustaining shifts across wakefulness and different stages of sleep.

This basic and well established set of assumptions has been further expanded by a series of reports claiming for a complementary role of dopaminergic mesolimbic activation and/or of dopaminergic mesocortical activation with the simultaneous absence of serotonergic cortical inhibitory modulation. The positive neurochemical feedback between the cholinergic and dopaminergic systems opens the question of the possible involvement of dopamine in sustaining the dreaming state of the brain. A raise in cortical dopamine activity has indeed been causally linked to the generation of nightmares, possibly implying an intensification of cholinergically-driven aspects of REM sleep-related dream mentation (for a detailed review on the neurochemical modulation of sleep, see Hobson et al., 2000). In summary, the periodicity of the presentation of single phases appears to be linked to feedback mechanisms between the different systems, on which the general homeostasis of the sleep-wake rhythm depends. The dramatic shifts in periodicity of the NREM-REM rhythm caused by the experimental selective deprivation of REM sleep appear to be consistent with this hypothesis.

The second intrinsic property of dreaming that can be found in acute psychosis is an impairment in conscious awareness. With the exception of so-called lucidity, the dreaming subject cannot consider the dream a mere construction of his own brain/mind until he is awake; the dreamer is not aware of dreaming but in contrast is fully immersed in a world experienced as real. Moreover, the dreaming subject loses the ability to critically evaluate what happens in the oneiric reality, so that the most improbable, bizarre and unreal things are experienced as natural. However, consciousness intended as the phenomenal level of the organization of the brain, with the sensory and spatial characteristics typical of the sense of personal presence in a multimodal reality, is fully realized in dreaming as in wakefulness (Revonsuo, 2000). The phenomenal first-person perspective is physiologically preserved with the sense of “me-ness” and of “phenomenal subjectivity”, meaning the awareness that the dynamics of experience rotate around a space centred on an experiential self (Metzinger, 2004).

Neurobiological substrates of dream consciousness supposedly lie in the functional interplay between cortical and subcortical structures. PET studies have shown that the REM stage of physiological sleep is characterized by a reduction in the metabolism of frontal structures, the dorsolateral prefrontal cortex (DLPFC) in particular, and a rise in metabolism of subcortical limbic structures and that dopaminergic mesolimbic reward circuits produce an instigating impulse for dreaming when they are activated by ascending brainstem arousal (Solms, 2000). This functional picture appears to be the basis of the subjective experience of

dreaming characterized by a relevant intensity in emotional charge and an almost total absence of metacognition and reflective awareness, functions linked to a normal activation of frontal cortical structures.

The third property of dreaming which appears to closely bind this form of mentation to psychosis is the presence of a complex sensory activity, predominantly visual and acoustic, of an hallucinatory nature. The full complexity of the physiological sensory activity of wakefulness can be found in the absence of external stimuli during dream sleep. Indeed, the dreaming brain/mind appears to be the biological proof that complex subjective experiences are possible in absence of sensorimotor activity: the dreaming world, though internally generated, is complete with its objects, spaces and actors which interact in often quite complicated events. Vision appears to be the central component of dreams, as dreaming most commonly means “seeing” (people, landscapes, places) and only in association to seeing can hearing or talking usually be experienced: in other words, whereas the hallucinatory visual experience is an essential condition of dreaming, hallucinatory experiences linked to the other sensory organs are optional.

The psychotic condition

A wealth of often contrasting definitions has been associated to the term psychosis since the early days of psychopathology. Little consensus has been reached to date, so that the Fourth edition of the Diagnostic and Statistical Manual opted for a restrictive definition of psychosis, to be preferably used as a term encompassing those serious mental disorders, including schizophrenia, major depression, alcohol withdrawal delirium, and others where the individual “loses touch with reality”. However, given the complexity and variety of clinical presentations, the Manual stresses the need to consider such a definition as an operative simplification (American Psychiatric Association, 2000). Karl Jaspers considered psychosis to be a lucid alienation encompassing bizarre behaviour, senseless furor, senselessness of affect and delusional ideas with a preserved orientation in space and time and a physiological level of vigilance; the loss of critical capacities and judgment over one’s experience is lost (Jaspers, 1963). Indeed, to be out of touch with reality means to present with signs of alienation accompanied by the absence of awareness of their being alien. Periodicity is an important property of this condition: acute psychosis presents clear waxing and waning within the individual life and the complexity of characterizing signs may disappear completely, or dampen and then reappear. The necessary and sufficient condition for the disappearance of signs to be clinically defined complete remission is for the subject to show a complete awareness of the impaired judgment of reality linked to the psychotic

mental state. The characterization of mental disorders presenting psychotic mental states in terms of periodicity is so important that their classification in periodic vs. non-periodic psychoses has been proposed.

Psychopathology and clinical psychiatry have indicated the full or partial retrieval of an awareness of reality as a differential element in the determination of the type of psychosis. From this viewpoint, for example, typical bipolar disorder is qualitatively different from typical schizophrenic disorder: the latter, though capable of yielding a silent psychopathology in certain phases with respect to typical signs (delusions, hallucinations or ideational incoherence) does never totally retrieve awareness of reality, as opposed to the former. Moreover, the time of life spent in a psychotic mental condition is directly correlated to a bad prognosis. This is true for mental disorders in which being psychotic is univocally inherent to the disorder itself (schizophrenia and paranoia – or delusional disorder – for example), but also when being psychotic is an optional state condition, as in affective disorders or substance abuse. On one side are subjects who show discrete psychotic periods with complete remission and on the other subjects presenting a psychotic state from which they never completely recover. It seems important to stress that this condition may arise in anybody: very seldom in subjects who are completely normal before and after the episode, more frequently in subjects affected with mental disorders that do not intrinsically imply an impaired judgment of reality, quite often in mental disorders characterized by mood oscillations and very frequently in two specific mental disorders: schizophrenia, where the psychotic condition is intermittent as far as productive symptoms are concerned (senseless furor, delusional ideas, bizarre behaviour, senselessness of affect) but continuous with regard to awareness of reality, and paranoia, where the psychotic condition is continuous in the two essential characteristics of delusional ideation and impaired judgment of reality.

Three strong lines of evidence support the hypothesis that the appearance of a psychotic mental state is supported by an increase in activity in the dopaminergic system: first, animal models that are considered closest to the psychotic mental state are induced by amphetamine and amphetamine-like substances; second, all substances which increase dopaminergic tone may precipitate acute psychoses; third, all substances which block post-synaptic dopamine receptors can effectively reduce signs of psychosis (Howes and Kapur, 2009). The increase in dopaminergic tone appears to be limited to mesolimbic pathways and accompanied by a contemporary fall in dopaminergic tone in the mesocortical pathways. This appears to lead to a dramatic fall in feedback control exercised by the DLPFC on limbic structures, which has been theoretically linked to the appearance of “positive” psychotic symptoms. The hypofrontality of the continuous psychotic mental state in schizophrenia has been confirmed by brain imaging techniques

(Minzenberg et al., 2009) and cognitive psychology has evidenced that the psychotic mental state is always accompanied by deficient performances in executive functions that are supposedly expression of dorsolateral prefrontal regions. The same cortical areas are thought to be involved in context analysis and in matching willed intention with acquired long-term memory patterns through a monitor system, and prefrontal cortex lesions and lesions of the cholinergic system ascending to the hippocampus have been found to destroy this complex neurofunctional pattern (Frith, 1991).

Hallucinatory phenomena involving different sensory modalities are often found in psychosis, with simple or complex visual experiences most commonly reported in organic conditions. "Voices" remain by far the most common type of hallucination in schizophrenia, although visual hallucinations are thought to be more frequent than generally reported. In the psychoses, internally generated percepts are usually experienced as sensory inputs from the external world, with no awareness of their pathological origin.

Connecting states of consciousness

The human brain has the ability to build a virtual world within which the Self finds itself immersed. This happens periodically during normal sleep and during wakefulness in the psychotic mental state. Both these conditions show extraordinary similarities in underlying cerebral functional patterns, with rises in cholinergic and dopaminergic tone in the neuronal pathways reaching the cortex and the limbic structures respectively from the brainstem. A fall in dopaminergic tone in the mesocortical pathways and a fall of the DLPFC activity is also necessary for both experiences to occur.

The common emerging picture is that of a broad activation of instinct and emotion-related structures with a simultaneous relative deactivation of inhibitory and executive cortical structures. This neurobiological framework is consistent with the original psychoanalytic model that highlights the instinctual origin of both dreaming and psychosis, with a consequent emergence of unconscious drives to consciousness. An obviously important difference between the dreaming and the psychotic brain/mind which currently finds no obvious explanation in terms of CNS functional physiology, is that psychosis coexists with varying levels of vigilance in an activated ARAS, whereas dreaming occurs within sleep, with the peculiar ARAS activation previously discussed. Indeed, psychosis occurs in both organic conditions such as delirium in which low or fluctuating levels of vigilance are found and so-called functional brain/mind states such as those found in schizophrenia or bipolar disorder in which vigilance is fully preserved. This

difference may account for the varying prevalence of different sensory modalities in the hallucinatory experience of dreaming and of different types of psychoses: conceptualized along a continuum, the brain's intrinsic ability to generate a multi-sensorial internal world is inversely correlated to the intensity of the external input. Interaction with the external world is primarily visual, so it is not surprising that on the end of the continuum in which this interaction is blocked internally generated inputs are most prominently visual. On the other end, when vigilance is fully maintained – as in mental disorders – the intensity of the external visual input is more likely to suppress the tendency to hallucinate visually and auditory hallucinations appear to be more prominent. Non-conscious brain processing, defined as a set of implicit operations in continuous development beneath the threshold of conscious awareness, may be hypothesized to represent the substrate of the internally generated input emerging when the subject's interaction with the external world is impaired. A possible explanation that appears to be consistent with this functional framework, is that of a mesolimbic dopaminergic hypertone responsible of the reduction of the inhibitory power of the GABAergic systems arising from the nucleus accumbens and reaching the cholinergic neural pathways of the forebrain, thus promoting psychosis. Interestingly, pharmacological compounds that are known to enhance GABA tone such as benzodiazepines have been hypothesized to have an antipsychotic efficacy, although recent investigations stress the need for new clinical data (Volz et al., 2007). On the other hand, in terms of sleep architecture, it is broadly accepted that these same agents suppress the occurrence of REM sleep.

In keeping with our hypothesis, both dreaming and psychosis may be considered physiological functioning modes of the brain/mind. Indeed, both conditions may be interpreted as the only recognizable mental activity when affective and emotional energy persists in absence of sensory inputs (dreaming), or within a chronic process of blockade of affective cathexis towards objects of the external world (psychosis). This view appears to be in striking agreement with Hughlings Jackson's hypothesis that the psychotic symptomatology found in alienated patients is what physiologically remains in a brain affected by an organic degenerative process. In these terms, schizophrenia or bipolar disorder may be viewed as clearly pathological, chronic mental conditions within which a dreamlike brain/mind activation is more likely to occur in wakefulness. Indeed, the continuum vs. dimension debate has recently began to shift in favour of several dimensions of psychosis occurring across diagnostic categories rather than discretely defined symptomatic entities. However, schizophrenia – a multifaceted disorder – is still the clinical condition within which psychosis may be diagnosed most frequently, and neurobiological and experiential correlates of psychosis are often derived from schizophrenic patients. For this reason, an erroneous equation schizophrenia =

psychosis is common in clinical practice and research settings, whereas several aspects of psychosis are known to occur in different disorders and schizophrenia with no relevant psychotic symptoms can occur. It is well known that a great variety of other predisposing conditions may precipitate psychosis, namely unipolar depression, mental retardation, drug intoxication or withdrawal and several metabolic or neurological conditions affecting the anatomy and physiology of the brain. Furthermore, psychosis can occur in otherwise normal subjects, in response to sensory deprivation or highly stressful events, and underlying neurobiological mechanisms have been hypothesized to involve stress-response cascades converging on dopamine pathways (van Winkel et al., 2008).

However, alternative to such determinism, a probabilistic view of brain functioning holds that the brain may express itself in an extraordinarily broad number of ways: the first concepts linked to this idea date back to the 1960s, when the physicist Walter Elsasser used the term *immense* to describe numbers greater than 10^{110} . The choice of this measure is not arbitrary, resulting from the mass of the Universe multiplied by its supposed age. The calculation of the “possible brains” in these terms yields a “hyperimmense” number, a number billions of times greater than 10^{110} (Scott, 1995): it appears, therefore, plausible to think of the psychotic mental state and of dreaming as two of the close to infinite ways of being of the human brain/mind.

Concluding remarks

The therapy of psychotic mental states is still largely unsatisfactory and mental health professionals struggle with the patients's loss of touch with reality. The dreamlike immersion in an internally generated world of perceptions and beliefs that are firmly held in contradiction with reality clearly pose a serious problem in the development of a therapeutic alliance. Emerging data from dream and sleep research may prove useful in psychiatric practice: the establishment of a conscious awareness of dreaming while still asleep is quite rare but has been shown to exist, and new quantitative EEG data seem to show that this mental state can be visualized in terms of a dissociated brain state (Voss et al., 2009). In this peculiar state of consciousness, termed “lucid dreaming”, the physiologically suppressed ego functions of dreaming appear to emerge either spontaneously in a limited number of individuals or through an induction technique requiring a standardized training.

Though data are still sparse, it has been reported that subjects most commonly become lucid by noticing bizarre elements within their dreams. Bizarreness

itself has long been recognized to represent a specific phenomenological aspect of dream cognition, although no consensus has been reached over its definition and several different theories have been developed to explain its occurrence: in our view, bizarreness in dreams may be best understood in terms of incongruities and discontinuities in the nature and mutual relationships amongst dream objects, thoughts and emotions. It seems plausible to hypothesize that the emergence of non-conscious processes results in the formally bizarre architecture of the conscious experience when the brain/mind's access to reality is blocked. The quantification of bizarreness is a useful experimental tool to compare the waking mentation of psychotic subjects with dream mentation, and an interesting overlap can be found (Scarone et al., 2008). Considering the dreaming subject's ability to control the dream experience once lucidity has been reached, the recognition of formally bizarre elements within one's own waking experience, be it in emotion, interpretation of reality, thought processes or hallucinatory phenomena, may prove of use in a subject's elaboration of psychosis. Studies into both the pharmacological and psychological induction of lucidity open a promising window for the comprehension of conscious/non-conscious interactions underlying the synthesis of what we call cognitive bizarreness and, consequently, the neurofunctional dynamic of psychotic mental states.

References

- American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders*, Fourth edition, text revision. Washington, DC: American Psychiatric Association.
- Freud, S. (1895). Project for a scientific psychology. In J. Strachey (Ed. & Trans.), *The standard edition of the complete works of Sigmund Freud – Volume 1* (1950). London, UK: Hogarth Press.
- Frith, C. (1991). In what context is latent inhibition relevant to the symptoms of schizophrenia? *Behavioral and Brain Sciences*, 14, 28–29.
- Hobson, J. A., Pace-Schott, E. F. and Stickgold, R. (2000). Dreaming and the brain: Toward a cognitive neuroscience of conscious states. *Behavioral and Brain Sciences*, 23, 793–842.
- Howes, O. D. and Kapur, S. (2009). The Dopamine Hypothesis of Schizophrenia: Version III—The Final Common Pathway. *Schizophrenia Bulletin*, 35, 549–562.
- Jaspers, K. (1963). *General Psychopathology*. Baltimore and London: Johns Hopkins University Press.
- Metzinger, T. (2004). *Being No One: The Self-Model Theory of Subjectivity*. Cambridge, MA: MIT Press.
- Minzenberg, M. J., Laird, A. R., Thelen, S., Carter, C. S. and Glahn, D. C. (2009). Meta-analysis of 41 functional neuroimaging studies of executive function in schizophrenia. *Archives of General Psychiatry*, 66, 811–822.
- Moruzzi, G. (1969). Sleep and instinctive behavior. *Archives Italienne de Biologie*, 107, 175–216.

- Nielsen, T. A. (2000). A review of mentation in REM and NREM sleep: "Covert" REM sleep as a possible reconciliation of two opposing models. *Behavioral and Brain Sciences*, 23, 851–866.
- Revonsuo, A. (2000). The Reinterpretation of Dreams: An evolutionary hypothesis of the function of dreaming. *Behavioral and Brain Sciences*, 23, 877–901.
- Scarone, S., Manzone, M. L., Gambini, O., Kantzas, I., Limosani, I., D'Agostino, A. and Hobson, J. A. (2008). The dream as model for psychosis: An experimental approach using bizarreness as a cognitive marker. *Schizophrenia Bulletin*, 34, 515–522.
- Scott, A. (1995). *Stairway to the Mind: The controversial New Science of Consciousness*. New York, NY: Springer-Verlag.
- Solms, M. (2000). Dreaming and REM sleep are controlled by different brain mechanisms. *Behavioral and Brain Sciences*, 23, 843–850.
- Taylor, E. (1983). *William James on exceptional mental states: The 1896 Lowell Lectures – Lecture 1*. New York, NY: Charles Scribner's Sons.
- van Winkel, R., Stefanis, N. C. and Myin-Germeys. (2008). Psychosocial stress and psychosis. A review of the neurobiological mechanisms and the evidence for gene-stress interaction. *Schizophrenia Bulletin*, 34, 1095–1105.
- Volz, A., Khorsand, V., Gillies, D. and Leucht, S. (2007). Benzodiazepines for schizophrenia. *Cochrane Database of Systematic Reviews*, Jan 24(1), CD006391.
- Voss, U., Holzmann, R., Tuin, I. and Hobson, J. A. (2009). Lucid dreaming: A state of consciousness with features of both waking and non-lucid dreaming. *Sleep*, 32, 1191–1200.

Conscious awareness versus optimistic beliefs in recreational Ecstasy/MDMA users

Andrew Parrott

The willingness of humans to take psychoactive drugs may reflect an unconscious optimism bias, where users focus on desired aims rather than actual consequences. A series of in-depth interviews will illustrate the experiences and explicit knowledge of recreational Ecstasy/MDMA users. Next an unpublished empirical study will be described, where four subgroups of Ecstasy users reported that MDMA loses its efficacy over time, while drug-related distress increased. As this cost-benefit ratio deteriorates, users take MDMA less frequently, before quitting permanently. The in-depth personal knowledge of experienced Ecstasy users might be useful for drugs education packages, since it could replace unconscious optimism with greater conscious awareness.

Lenton, Boys and Norcross (1997) investigated the drugs-knowledge of 83 rave attendees in Australia, many of whom had first taken Ecstasy/MDMA while still at school. They noted that: 'Many new ravers were found to have romantic positive beliefs about the effects of Ecstasy, LSD and speed, and were effectively blind to the more negative effects'. They also found that novice drug users had significantly less knowledge than experienced users. Romantic optimism about recreational drugs is widespread. It is often present in the web-pages designed for recreational drug users. The Internet site 'Dancesafe' has attractive butterflies flying across its Ecstasy pages, and this may unconsciously foster pro-Ecstasy beliefs. The verbal information also focuses more on acute mood gains, rather than longer-term problems. The name 'Dancesafe' reflects this optimism bias, and does not acknowledge that recreational stimulants are *inherently* damaging, irrespective of how they are taken (Parrott, 2001; 2006). Many experienced users do however display extensive drugs knowledge, with surveys revealing numerous Ecstasy-attributed problems (Parrott et al., 2002; Rodgers et al., 2003; Topp et al., 1999). Some recreational users also argue for 'informed choice', stating that decisions about drug-taking should be conscious, reflective, and well-informed. There is little evidence to support this model among young non-users, who are

mainly 'at risk' from unconscious social pressures. The aim of the chapter is to examine the knowledge of Ecstasy/MDMA users, to see if it might be useful for drugs education packages. Their personal experiences could increase conscious awareness in younger 'at risk' individuals. It might replace unconscious optimistic beliefs, with more realistic information about the practical consequences of becoming a drug user.

Interviews with Ecstasy/MDMA users

Each year, the undergraduate students on my 'Drugs and Behaviour' course at Swansea University conduct detailed interviews with drug users. They devise their own semi-structured questionnaires, covering topics such as positive and negative experiences, first usage, recovery phenomena, and longer-term effects. In our most recent cohort, five students conducted interviews with Ecstasy/MDMA users whom they knew. Their reports are summarized in Table 1. The on-drug experiences were described in very positive terms: 'love everyone', 'very euphoric and a feeling of closeness', 'nothing as good – that is why they call it Ecstasy' (Table 1). These findings were similar to previous studies (Cohen, 1998; Parrott, 2001), and reflect the increased levels of serotonin, dopamine, noradrenaline and other neurotransmitters (Green et al., 2003). These neurotransmitters underlie many different psychobiological functions, so that numerous side-effects can occur. These includes jaw-clenching, gurning, and tooth-grinding; hence: 'chewing mouth and lips all night so they can get very swollen' (Table 1), with long-term dental problems which can include tooth erosion. MDMA also affects the hypothalamus, and impairs temperature regulation and water intake (Green et al., 2003). Morefield et al. (2009) found an average temperature increase of $+0.8^{\circ}\text{C}$ in recreational MDMA users, hence the current reports of 'feeling hot' (Table 1). For several days following MDMA, the body is in a state of neuropsychobiological recovery, with low energy, impaired sleep, reduced appetite, and feelings of paranoia (Table 1). Again, this agrees with previous reports of mid-week problems including, lethargy, depression, anger, disrupted sleep, and acute psychiatric distress (Cohen, 1998; Jones et al., 2008; Parrott and Lasky, 1998; Schifano, 2000; Verheyden et al., 2003). The long-term effects of regular MDMA use can include poor memory, reduced social intelligence, sleep apnoea, weight loss, impaired immunosuppression, sexual dysfunction, and various psychiatric symptoms (Parrott, 2001; 2006; Rodgers et al., 2003; Schifano, 2000; Topp et al., 1999).

Table 1. Interviews with five ecstasy/MDMA users**Female aged 21, taking Ecstasy/MDMA for 2 years.**

First usage: 'I was offered them at a club by my friend when I was drunk. It was a split second decision and I didn't really know what it was I was taking, but had I had more information and been sober, I would still have tried them that night.'

Continued usage: 'If I used pills now I would take 3 or 4, but I've taken more than this in the past. When I first used them I wouldn't have taken more than 2. I wouldn't take more than ½ gm of MDMA now – I think more than ½ gm can be dangerous.'

On-drug effects: 'very euphoric and a feeling of closeness', 'mildly hallucinogenic'.

Off-drug recovery period: 'no energy/motivation to do anything', 'can be impossible to sleep', 'appetite massively affected the first few occasions', 'either no mood change, or a negative one which could range from mild to severe', 'lot of inward thinking'. Sensitivity increased: 'almost cried because I could not find something I wanted in the supermarket', feelings of 'mild paranoia', and 'once had extreme paranoia when I got home after taking pills. I had heart palpitations and was very panicky...scared to leave my room'.

Cessation: Still uses occasionally, but frequency of usage had reduced. 'occasionally I have been sick of pills and MDMA'.

Male former-user aged 24; had used when 18–21 years old.

First usage: 'began using MDMA with a group of friends during a dark and depressive phase of his life. He does not go into much detail about this'.

Continued usage: Usually took ½ pill and did not increase this for a long time. With larger supplies he repeated its usage. MDMA powders 'more expensive to acquire and much more intensive in effect'.

On-drug effects: 'more sociable and amiable', 'very embarrassed at memories exemplifying disinhibition when on drug... expressly dislikes dancing and remembers doing so with distaste', 'more active sex life... qualitatively less successful', 'stronger desire to drink water paired with cautiousness due to a rumour of being able to drown oneself', 'feeling very hot or very cold even if the actual temperature did not match', 'twitching or shaking of the leg and continuous rolling of cigarettes... use the rolling habit to stock up on cigarette supply'.

Off-drug recovery: 'one or two days of bad moods and tiredness', 'repressed appetite, feeling extremely fatigued and being generally disinterested', 'sleep pattern very irregular'.

Cessation: had now quit using. 'He could not see himself functioning appropriately in today's society while continuing drug use'. 'Too strongly negative effects upon his appetite'. 'Decided to quit one day and did so without any problems of withdrawal'.

Female aged 19 years, using MDMA for 8 months.

First usage: 'wanted to try it', 'have a mind set to try everything once'

Tolerance: 'only used to take one bomb a night'...now takes: 'three or more bombs a night' [Note: a bomb is MDMA powders in a twist of cigarette paper, which is swallowed], 'would never pay for drugs before but now pay for MDMA'.

On-drug effects: 'more aware. Loads of energy to dance, buzzing. Love everyone. Giggly'.

Off-drug recovery: 'body feels heavy. Need to relax, don't want to do anything. Sometimes feel hot and sweaty and have a headache. Just want to sleep through it'. Housemates say she should not take it as 'they do not like it', but 'they are ignorant to drugs, don't know anything'.

Cessation: 'Only take it on certain nights, can go for months without it'

Male 20 years old. Duration of usage and amounts not described.

First usage: 'influenced by friends at school. When they did it he followed suit.'

Continued usage: 'to avoid this state of dependency and having to constantly increase the dose to achieve the same hit, he varies how often he takes MDMA, and now leaves 'a few months' between uses.

On-drug effects: 'nothing like it, nothing as good – that why they call it ecstasy', 'feel really energetic... want to jump around and dance', 'nerve endings are alive and just tingling all over the body', 'chews his mouth and lips all night so they can get very swollen'.

Post-drug recovery: 'bad come down', 'impossible to sleep', 'Don't want to eat at all... if continually took it you wouldn't need to eat for a week', 'after taking a bit hit one can be very vague and forgetful for about a week but then your memory goes back to normal'.

Cessation: 'Doesn't think it is addictive' 'Feeling when you are out sometimes that you want some when out clubbing, but its not a need, its just a indulgence.' 'It is better to take it when you are younger, because when people try it in their early twenties they are more like to start off doing more and you have more at stake to lose, for example their university degree'.

Male novice user 21 years, taken around 4–5 times.

First usage: 'first tried MDMA at a house party 5 months ago. His friends were using it and out of curiosity he tried it to see what it was like. He said he lives by the mantra 'try everything once' so decided to give it a go after his friends had explained the pros and cons of the drug'.

Continued usage: 'first tried a small amount', subsequently using 'around 500 mg per trip'.

On-drug effects: 'imagine the best feeling you have ever felt, times it by ten and you're still not close to how amazing you feel', 'cold things feel great', 'salt and vinegar crisps taste disgusting', 'looking at patterns is mind-boggling. I once thought a plastic fish was swimming on the spot out of water'.

Off-drug recovery: 'Short term memory loss. Sometimes you will be talking to someone and mid sentence you will forget what you wanted to say', 'difficulty sleeping'.

Cessation: 'Had never felt an urge or craving to take MDMA' 'If he did feel a craving it would be more for the social aspect than the physiological'.

With regard to drug initiation, the main reasons given were peer influence and curiosity: 'Friends were using', 'To see what it was like', and 'Influenced by his friends at school' (Table 1). This is typical of all recreational drugs, whether legal or illicit (Parrott et al., 2004). Dancing and partying were also noted, which is consistent with Ecstasy/MDMA being a dance club drug. Informed choice was only noted by one of the five interviewees, who 'decided to give it a go after his friends had explained the pros and cons of the drug.' The interview report did not describe what these friends had told him, although in previous interviews with Ecstasy users, the information typically comprised practical advice, such as drink fluids regularly, and rest when you overheat. Indeed many young drug initiates seem to be aware of these types of practical guideline (Lenton et al., 1997).

However, there was no evidence that any novice user displayed any real understanding of how and why MDMA was inherently damaging (Table 1). Hence there was minimal support for the 'informed choice' model. Indeed, in fifteen

years of Ecstasy/MDMA research, I have come across only one young person who had undertaken any *scientific reading* prior to taking MDMA. This was an undergraduate student who had read about the neurotoxic effects of MDMA in a popular science journal. Then when his friends started to take ecstasy, he argued against them, stating that it was not safe. He regularly went clubbing with them – but without taking the drug himself. Some time later he read *another* scientific article about the USA government giving MDMA a license for psychotherapy, and so he decided to take it. That was two years previously, and now he was a volunteer in one of our laboratory studies. Subjectively he complained of various problems including memory deficits, and his cognitive performance on some of the laboratory tasks was impaired.

Two of the five users stated that they took Ecstasy because they wanted to ‘try everything once’ (Table 1). However neither of them stuck with this intention, and instead had taken it on multiple occasions. Furthermore, all five users had increased their dosage levels over time (Table 1). Most users start with a single tablet, before increasing the number of tablets per occasion. MDMA powders have recently become more popular, and they provide a further means for taking higher doses, since MDMA ‘bombs’ can contain the equivalent of 2 or 3 Ecstasy tablets (150–225 mg MDMA). The trend for higher doses reflects chronic tolerance, or the reduction in drug efficacy with repeated usage, and reflects neuro-adaptive reactions in the brain and liver (Parrott, 2005). The rate of development of chronic tolerance can be variable, with two of the interviewees reporting rapid dosage escalation, but another showing minimal escalation (Table 1). An Australian study of careful non-ravers also showed slower dosage escalation, and interviewee 2 also seems to have not been much of a raver: ‘Expressly dislikes dancing and remembers doing so with distaste’ (Table 1). Environmental co-stimulation may indeed exacerbate the neuropsychobiological effects of MDMA (Parrott, 2005; 2006). Few drug users are however consciously aware of drug tolerance. Instead they often attribute their need for higher doses as due other factors; the most common being Ecstasy tablets are ‘getting weaker’. This is similar to alcohol drinkers complaining about beer being, ‘watered-down’ [note: the empirical evidence shows similar Ecstasy/MDMA dosage levels over time].

One intriguing aspect of Ecstasy/MDMA is the variation in cognitive awareness between individuals, and differences in recognition and admission of drug-related problems. Fox et al. (2001) compared the cognitive profiles of Ecstasy users who said that *did* have problems due to Ecstasy/MDMA, with a subgroup who stated that they *did not* have problems. Performance levels on two cognitive planning and spatial memory tasks were significantly impaired in both groups. These cognitive deficits were significantly linked to lifetime usage, with more

problems in the heavy Ecstasy/MDMA users, irrespective of either conscious awareness, or explicit recognition. Variation in problem-recognition also occurred here (Table 1). One interviewee noted that her flatmates did not like her use of MDMA, but stated that their comments were not valid since 'they are ignorant to drugs, don't know anything'. This illustrates the cognitive rationalization that often surrounds drug usage. Others displayed more awareness, with a former user noting: 'He could not see himself functioning appropriately in today's society while continuing drug use' (Table 1). An interesting topic for future research is meta-cognition. How do different users attribute their problems? How important are unconscious processes such as denial and repression – do they facilitate continued usage?

Awareness of the positive and negative effects of Ecstasy/MDMA

In this second study, we investigated the explicit conscious awareness of positive and negative drug effects. Four subgroups of current and former ecstasy/MDMA users completed an Ecstasy/MDMA effects questionnaire (Parrott and O'Sullivan, unpublished). Table 2 presents the findings from four questions: positive effects on the first occasion they had taken MDMA, negative effects on the first occasions, and these same two questions for their most recent Ecstasy/MDMA experience. Light users who had taken Ecstasy/MDMA a few times before quitting, showed lower 'positive' drug experiences on their first occasion, than the other three groups (Table 2). Their less euphoric drug reaction, may provide a reason for their quitting usage after just a few experiences. It agrees with Louisa Degenhardt, who has noted that many who experiment with drugs find the effects aversive and do not persist. Hence not every reaction to MDMA is favorable, with marked variance in subjective mood reactions. Hence many reactions to MDMA comprise a mixture of positive and negative changes (reviews: Parrott, 2001; 2006).

Another general finding here was that positive effects declined from the first to the last drug experience, while negative effects increased (Table 2). This was consistent with some early findings by Peroutka (cited in Parrott, 2005). Every subgroup showed these changes, although they occurred to different extents. The two heavy user subgroups, who had taken most drug overall, showed the greatest declines in mood (Table 2). This illustrates the chronic tolerance noted earlier. The positive and negative self-report data were used to calculate cost-benefit ratios (See Chapter 15 in Parrott et al. 2004, for a debate about cost-benefit ratios). All four groups reported beneficial cost-benefit ratios the first time they took MDMA, although the score for novice-quitters was noticeably lower (Table 1). Every group also reported a decline in cost-benefit ratio with repeated drug usage.

The only group to still report a positive ratio – with more gains than losses, was current novice users. The ratios for both groups of experienced users were negative (Table 2). Furthermore, the frequency of MDMA use tends to decline over time (Verheyden et al., 2002), a pattern of usage that is almost unique amongst recreational drugs (Parrott, 2005). In other informal interviews, experienced Ecstasy/MDMA have told us that as their negative reactions increase, they take the drug less frequently, before quitting permanently.

Table 2. Positive and negative effects for the first and last times on Ecstasy. Self-reported questionnaire scores for four subgroups of recreational users: divided into light and heavy lifetime use, current and former users (Unpublished study: Parrott & O'Sullivan, unpublished)

Dependent variable	Light former users	Light current users	Heavy former users	Heavy current users	ANOVA Group Effect F-value (df 3,60) & significance level
Sample size (N)	18	10	22	18	—
Positive effects <i>first time</i>	4.9	6.6	8.6	8.8	F = 12.22, p < 0.001
Positive effects <i>last time</i>	3.9	6.2	3.4	4.3	F = 4.40, p < 0.01
Decline in positive effects: – from <i>first to last time</i>	–1.0	–0.4	–5.2	–4.5	—
Negative effects <i>first time</i>	3.6	3.2	3.9	2.9	F = 0.55, non-sig
Negative effects <i>last time</i>	5.8	4.2	5.3	5.0	F = 0.66, non-sig
Increase in negative effects: – from <i>first to last time</i>	+2.2	+1.0	+1.4	+2.1	—
Cost-Benefit ratio <i>first time</i>	+36%	+103%	+120%	+203%	—
Cost-Benefit ratio <i>last time</i>	–33%	+47%	–36%	–14%	—

Self-rating scales: 1–10 (minimum–maximum).

Conclusions

Most young people have little conscious knowledge about drugs. Instead many of their beliefs are unrealistically optimistic, reflecting unconscious social influences (Lenton et al., 1998). The aim of this chapter is to discuss how their low levels of conscious awareness might be improved. MDMA is a powerful CNS stimulant (Green et al., 2003), which adversely affects core psychobiological systems for sleep, memory, cognition, immunocompetence, and psychiatric well-being (Parrott, 2001; 2006; Schifano, 2000; Jones et al., 2008; Topp et al., 1999); some of

these adverse effects were illustrated here (Table 1). Furthermore, the on-drug experiences of Ecstasy users decline with repeated usage (Table 2). This information could be very useful for drugs education packages. The personal stories of experienced users, and their reasons for quitting, might be something which a younger cohort would relate to. This personal knowledge might also be supplemented by basic scientific findings. In a recent study of recreational Ecstasy users, MDMA use while dancing led to an 800% increase in the stress hormone cortisol (Parrott et al., 2008). In post-study feedback sessions, the Ecstasy users were impressed by the strength of this effect, and particularly interested in how this stress neurohormone is often associated with psychobiological problems (Parrott, 2009). There is also evidence for serotonergic neurotoxicity in regular users (Cowan, 2007), and altered brain microvasculature in novice users (de Win et al., 2008). Given this wide range the problems, it is difficult to see how anyone with a limited knowledge of its adverse effects on the human body and brain, would decide to take it. If the 'informed choice' model were correct, then we would have very few recreational MDMA users.

References

- Cohen, R. S. (1998). *The Love Drug: Marching to the Beat of Ecstasy*. New York State: Haworth Medical Press.
- Cowan, R. L. (2007). Neuroimaging research in human MDMA users: A review. *Psychopharmacology*, 189, 539–556.
- de Win, M. M., Jager, G., Booij, J., Reneman, L., Schilt, T., Lavini, C., Olabarriaga, S. D., den Heeten, G. J. and van den Brink, W. (2008). Sustained effects of ecstasy on the human brain: a prospective neuroimaging study in novel users. *Brain*, 131, 2936–2945.
- Fox, H., Parrott, A. C. and Turner, J. J. D. (2001). Ecstasy/MDMA related cognitive deficits: A function of dosage rather than awareness of problems. *Journal of Psychopharmacology*, 15, 273–281.
- Green, A. R., Mechan, A. O., Elliott, J. M., O'Shea, E. and Colado, M. I. (2003). The pharmacology and clinical pharmacology of 3,4-methylenedioxymethamphetamine (MDMA, "ecstasy"). *Pharmacology Reviews*, 55, 463–508.
- Jones, K. A., Callan, F., Blagrove, M. T. and Parrott, A. C. (2008). Sleep, energy and self-rated cognition across 7 nights following recreational ecstasy/MDMA use. *Sleep and Hypnosis*, 10, 16–28.
- Lenton, S., Boys, A. and Norcross, K. (1997). Raves, drugs and experience: Drug use by a sample of people who attend raves in Western Australia. *Addiction*, 92, 1327–1337.
- Morefield, K. M., Keane, M., Felgate, P., White, J. M. and Irvine, R. J. (2009). The acute psychobiological impacts of illicit 3,4-methylenedioxymethamphetamine (MDMA, 'Ecstasy') consumption in recreational environments. *Neuropsychobiology*, 60, 213–223.
- Parrott, A. C. (2001). Human psychopharmacology of Ecstasy (MDMA): A review of fifteen years of empirical research. *Human Psychopharmacology*, 16, 557–577.

- Parrott, A. C. (2005). Chronic tolerance to recreational MDMA (3,4-methylenedioxymethamphetamine) or Ecstasy. *Journal of Psychopharmacology*, 19, 71–83.
- Parrott, A. C. (2006). MDMA in humans: Factors which affect the neuropsychobiological profiles of recreational Ecstasy users, the integrative role of bio-energetic stress. *Journal of Psychopharmacology*, 20, 147–163.
- Parrott, A. C. (2009). Cortisol and MDMA (3,4-methylenedioxymethamphetamine): Neurohormonal aspects of bioenergetic-stress in Ecstasy users. *Neuropsychobiology*, 60, 148–158.
- Parrott, A. C. and Lasky, J. (1998). Ecstasy (MDMA) effects upon mood and cognition; before, during, and after a Saturday night dance. *Psychopharmacology*, 139, 261–268.
- Parrott, A. C., Buchanan, T., Scholey, A. B., Heffernan, T. M., Ling, J. and Rodgers, J. (2002). Ecstasy attributed problems reported by novice, moderate and heavy recreational users. *Human Psychopharmacology*, 17, 309–312.
- Parrott, A. C., Morinan, A., Moss, M. and Scholey, A. (2004). *Understanding Drugs and Behaviour*. Chichester: Wiley.
- Parrott, A. C., Lock, J., Conner, A. C., Kissling, C. and Thome, J. (2008). Dance clubbing on MDMA and during abstinence from Ecstasy/MDMA: Prospective neuroendocrine and psychobiological changes. *Neuropsychobiology*, 57, 165–180.
- Rodgers, J., Buchanan, T., Scholey, A. B., Heffernan, T. M., Ling, J. and Parrott, A. C. (2003). Patterns of drug use and the influence of gender on self reports of memory ability in ecstasy users: A web based study. *Journal of Psychopharmacology*, 17, 379–386.
- Schifano, F. (2000). Potential human neurotoxicity of MDMA ('Ecstasy') subjective self-reports, evidence from an Italian drug addiction centre and clinical case studies. *Neuropsychobiology*, 42, 25–33.
- Topp, L., Hando, J., Dillon, P., Roche, A. and Solowij, N. (1999). Ecstasy use in Australia: Patterns of use and associated harm. *Drug Alcohol Dependence*, 55, 105–115.
- Verheyden, S. L., Henry, J. A. and Curran, H. V. (2003). Acute, sub-acute and long-term subjective consequences of 'ecstasy' (MDMA) consumption in 430 regular users. *Human Psychopharmacology*, 18, 507–517.

Conscious and unconscious placebo responses

How the ritual of the therapeutic act changes the patient's brain

Fabrizio Benedetti

The study of the placebo response is basically the study of the psychosocial context around the therapy, which constitutes the ritual of the therapeutic act, and of its effects on the patient's brain. Many mechanisms are involved, both conscious, like expectation of a future outcome, and unconscious, such as classical conditioning. Overall, recent research indicates that different social stimuli, first and foremost the therapist's words, may induce cellular and molecular changes in the patient's brain, thus placing psychotherapy into the neurobiological domain.

The placebo effect, or response, has recently emerged as a model to understand how the brain influences, both consciously and unconsciously, the course of disease and the response to a therapy. Until a few years ago the placebo effect was considered a nuisance in clinical research whenever the efficacy of a new therapy had to be validated. In a typical clinical trial, the treatment under test is compared against a placebo, that is, an inert treatment that has no specific effect for the pathology that has to be treated. For example, in a clinical trial for the assessment of a new painkiller, a new pharmacological agent is compared against a saline solution (the placebo) which, of course, has no analgesic properties. If the pharmacological agent relieves pain better than the placebo, it is considered to be an effective painkiller. Conversely, if the drug and the placebo produce the same therapeutic outcomes, the drug is considered to be ineffective.

In a pragmatic clinical trial, clinicians and clinician scientists are usually not interested to understand why the patients who receive an inert treatment (the placebo) show a clinical improvement. What matters is that the therapy under test be better than the placebo. In contrast to this pragmatic approach, which is

only aimed at assessing whether or not a therapy works, in recent years there has been a growing interest in understanding why the patients who get the placebo improve. Needless to say, neither water solutions nor sugar pills will ever acquire therapeutic properties. What matters is the psychosocial context within which the placebo is administered. In other words, the study of the placebo effect is basically the study of the psychosocial context around the patient and the treatment and how these social stimuli affect the patient's brain. Anything that surrounds the therapy is part of the psychosocial context, for example, doctors, nurses, syringes, complex medical machines and, probably most important, the words by the therapist. All these social/sensory stimuli tell the patient that a therapy is being performed and, as a whole, they represent the so-called therapeutic act (Benedetti, 2008b).

The idea of a top-down control on several physiological processes by a variety of psychological and social factors is not new, of course. Indeed, the top-down modulation of sensory inputs plays a fundamental role in shaping the global perceptual experience. What is somehow different for the placebo effect is that this top-down cognitive and emotional modulation occurs in the clinical setting, whereby the intensity and severity of symptoms can be shaped by the psychological state of the patient and by the social context surrounding the treatment (Benedetti, 2008a, b; Price et al., 2008). Therefore, it goes without saying that this has profound implications for clinical practice, pharmacotherapy and psychotherapy, as well as for our understanding of how the brain controls different physiological processes throughout the body.

What is emerging today is that the psychosocial context around the therapy, namely the therapeutic act itself, activates several molecular and cellular mechanisms in the patient's brain that are the same as those activated by drugs (Colloca and Benedetti, 2005). In other words, social stimuli and drugs may activate common biochemical pathways, and this occur through both conscious and unconscious mechanisms.

Conscious placebo responses

By definition, a placebo effect is the effect that follows the administration of a placebo, that is, of an inert treatment. As the inert treatment is given along with contextual stimuli, for example verbal suggestions of clinical improvement which make the patient believe that the treatment is real and effective, its study is basically the study of the context around the treatment and how it affects the patient's brain. When the context is positive and leads to positive outcomes, the term placebo effect is used. Conversely, when the context is negative and leads to negative outcomes,

the term *nocebo effect* is adopted. Therefore, the placebo effect is a psychobiological phenomenon, i.e. something active happening in the patient's brain. It has nothing to do with spontaneous remission, patients' and/or physicians' biases, the effects of unidentified co-interventions, and many other factors, which are all phenomena that are sometimes mistakenly taken for placebo effects.

There is not a single mechanism of the placebo effect and there is not a single placebo effect but many, so that we have to look for different mechanisms in different medical conditions and in different therapeutic interventions (Benedetti, 2008a, b). Most of the research on placebos has focused on expectations as the main factor involved in placebo responsiveness. Indeed, the literature is full of studies whereby expectations are analyzed, and the terms "effects of placebos" and "effects of expectations" are frequently used interchangeably. In general, the expectations of a future outcome and of a future response, the so-called response expectancies, are held by an individual about one's own emotional and physiological responses such as pain, anxiety and sexual arousal (Kirsch, 1999). Expectations may lead to positive or negative outcomes in a number of ways. From a neuroscientific point of view, expecting a future event may involve several brain mechanisms that are aimed at preparing the body to anticipate that event. For example, the expectation of a future positive outcome may reduce anxiety and/or activate the neuronal networks of reward mechanisms, whereas the expectation of a negative outcome is aimed at anticipating a possible threat, thus increasing anxiety (Enck et al., 2008). Indeed, anxiety has been found to be reduced after placebo administration in some studies, for example in patients with irritable bowel syndrome (Vase et al., 2005).

That anxiety is an important mediator of placebo responses is indicated by the opposite effects, that is, the *nocebo* responses. The *nocebo* effect is a phenomenon that is opposite to the placebo effect, whereby expectation of a negative outcome may lead to the worsening of a symptom. It basically consists in delivering verbal suggestions of negative outcomes so that the subject expects clinical worsening. Recent experimental evidence indicates that negative verbal suggestions induce anticipatory anxiety about the impending pain increase, and this verbally-induced anxiety triggers the activation of cholecystokinin (CCK) which, in turn, facilitates pain transmission. In fact, CCK-antagonists have been found to block this anxiety-induced hyperalgesia (Benedetti et al., 2006).

Expectations may also induce changes through reward mechanisms, which ensure future reward acquisition (Enck et al., 2008). These mechanisms are mediated by specific neuronal circuits linking cognitive, emotional, and motor responses, and are traditionally studied in the context of the pursuit of natural (e.g., food), monetary, and drug reward. Within the context of placebo responses, whereby the reward is the therapeutic benefit and the clinical improvement

itself, de la Fuente-Fernandez et al. (2001), by using positron emission tomography, found dopamine activation in the nucleus accumbens, a major region of the reward circuitry, following the administration of a placebo in patients with Parkinson's disease. Interestingly, no relationship was found between nucleus accumbens dopamine activity and the actual placebo effects on motor function, thus suggesting that dopamine activation was better related to the expectation of reward. Likewise, Mayberg et al. (2002) observed nucleus accumbens activation in depressed patients after one week of placebo treatment, and Scott et al. (2007) found a correlation between the responsiveness to placebo and that to monetary reward, along with the activation of dopamine in the nucleus accumbens. Scott et al. (2008) also found an activation of opioid receptors in the nucleus accumbens during placebo analgesia, thereby confirming previous pharmacological and imaging studies on the involvement of endogenous opioids in placebo analgesia (Amanzio and Benedetti, 1999; Wager et al., 2007). Therefore, placebos activate the same receptorial pathways that are activated by drugs, namely, dopamine receptors are activated by both placebos and anti-Parkinson drugs and, similarly, opioid receptors are activated by both placebos and opioid analgesics.

Learning is another mechanism that is central to placebo responsiveness. Subjects who suffer from a painful condition, such as headache, and who regularly consume aspirin, can associate the shape, colour and taste of the pill to pain decrease. After repeated associations, if they are given a sugar pill that looks like aspirin, pain can decrease. In place of the shape, colour and taste of pills, several other stimuli can be associated to clinical improvement, such as syringes, stethoscopes, white coats, hospitals, doctors, nurses, and so on. The mechanism that underlies this effect is conditioning, whereby a conditioned (neutral) stimulus, e.g. the colour and shape of a pill, can become effective in inducing the reduction of a symptom if repeatedly associated to an unconditioned stimulus, i.e. the drug inside the pill. This type of associative learning may represent the basis of many placebo effects, whereby the placebo is the conditioned (neutral) stimulus itself. In this regard, it is fundamental to point out that several forms of conditioning include important cognitive components. In fact, many authors dissent from the view that conditioning is an automatic unconscious event that is due to the temporal contiguity between the conditioned and the unconditioned stimulus. Indeed, conditioning has been reinterpreted in cognitive terms on the basis that conditioned learning does not depend simply on the pairing of the conditioned and the unconditioned stimuli, but on the information that is contained in the conditioned stimulus (Rescorla, 1988). In other words, conditioning would lead to the expectation that a given event will follow another event, and this occurs on the basis of the information that the conditioned stimulus provides about the unconditioned stimulus (Rescorla, 1988).

That cognition plays a crucial role in placebo analgesia has recently been shown in a study in which social learning was found to be as effective as conditioning in inducing robust placebo responses. Colloca and Benedetti (2009) compared placebo analgesia induced through social observation with first-hand experience via a typical conditioning procedure and with verbal suggestion alone. In the social condition, subjects underwent painful stimuli and placebo treatment after they had observed a demonstrator (actually a simulator) showing analgesic effect when the painful stimuli were paired to a green light. In the conditioning experimental condition, subjects were conditioned according to a classical conditioning procedure, whereby a green light was associated to the surreptitious reduction of stimulus intensity, so as to make them believe that the treatment worked. In the verbal suggestion condition, subjects received painful stimuli and were verbally instructed to expect a benefit from a green light. It was found that observing the beneficial effects in the demonstrator induced substantial placebo analgesic responses which were positively correlated with empathy scores. Moreover, observational social learning produced placebo responses that were similar to those induced by directly experiencing the benefit through the conditioning procedure, whereas verbal suggestions alone produced significantly smaller effects. Thus, social observation is as powerful as conditioning in producing substantial placebo responses, which indicates that conditioned placebo analgesia is not only an automatic process, but it involves cognitive factors.

Unconscious placebo responses

The ritual of the therapeutic act can affect the patient's brain through totally unconscious processes. In other words, one needs not believe in the treatment and trust his doctor in order to respond to a placebo treatment. Immune and hormonal responses represent two good examples of unconscious placebo responses, whereby the underlying mechanism is likely to be classical, or Pavlovian, conditioning. This aspect of placebo responsiveness is important because conditioned placebo responses take place regardless of what the patient expects. For example, conditioned hormonal increases can be obtained even if the subject expects hormonal decreases (Benedetti et al., 2003).

Some of the first compelling evidences that immunological responses can be behaviourally conditioned was obtained by a long series of experiments performed in the 1970s and 1980s. By using a taste aversion conditioning paradigm in rats, Ader and Cohen (1975) paired a flavoured drinking solution (saccharin) with the immunosuppressive drug, cyclophosphamide. The rats were subsequently immunized with sheep red blood cells. Those rats that were re-exposed to saccharin at

the time of antigenic stimulation were found to have lower hemagglutinating antibodies six days after the injection of the sheep red blood cells compared to conditioned animals that were not re-exposed to saccharin, non-conditioned animals given saccharin and a placebo group. Thus, saccharin was capable of mimicking the immunosuppressive action of cyclophosphamide. Many subsequent studies confirmed these findings in a variety of experimental models.

What is more interesting is perhaps the finding that conditioned immune responses can also be obtained in humans. Although some early studies produced contrasting results, more recent research provides convincing evidence that behavioural conditioning of immunosuppression is possible in humans (Goebel et al., 2002). Repeated associations between cyclosporin A and a flavoured drink induced conditioned immunosuppression in healthy male volunteers, in which the flavoured drink alone produced a suppression of the immune functions, as assessed by means of interleukin-2 (IL-2) and interferon- γ (IFN- γ) mRNA expression, *in vitro* release of IL-2 and IFN- γ , as well as lymphocyte proliferation. It is interesting to note that the effects of the conditioned stimulus were virtually the same as those of the specific effects of cyclosporine A.

In the endocrine system, similar effects can be found. For example, hypoglycaemia can be conditioned in humans by pairing insulin with a conditioned stimulus (Stockhorst et al., 1999). In order to assess whether a cognitive component is present in hormonal conditioning, the study by Benedetti et al. (2003) was aimed at differentiating the effects of conditioning and expectation on plasma levels of growth hormone and cortisol. In the first experimental condition, verbal suggestions of growth hormone increase and cortisol decrease were delivered to healthy volunteers, so as to make them expect hormonal changes. These verbal instructions did not have any effect on both hormones, and in fact no plasma concentration change was detected. In the second experimental condition, sumatriptan, a serotonin 5-HT_{1B/1D} receptor agonist that stimulates growth hormone and inhibits cortisol secretion, was administered for two days in a row and then replaced with a placebo on the third day. A significant increase of growth hormone and decrease of cortisol plasma concentrations were found after placebo administration. These conditioned effects occurred regardless of the verbal suggestions the subjects received. In other words, the placebo mimicked the sumatriptan-induced growth hormone increase, even though the subjects expected a growth hormone decrease. Likewise, the placebo mimicked the sumatriptan-induced cortisol decrease, even though the subjects expected a cortisol increase. It can be assumed that in this case the conditioned stimulus was represented by the act of injecting the pharmacological agent (i.e., the context around the treatment). The experiment by Benedetti et al. (2003) clearly shows that sometimes it is not necessary to expect anything in order for a placebo response to occur.

These represent the best examples of unconscious placebo effects that take place even in the absence of the patient's expectations and beliefs.

When do conscious and unconscious placebo responses occur?

The effects of the therapeutic act on the patient's brain are attributable to both conscious expectation mechanisms and unconscious conditioning mechanisms. In some circumstances, expectation can override a conditioning procedure, whereas in some other conditions, conditioning overrides conscious expectations. In this regard, the study by Benedetti et al. (2003) was aimed at inducing conflicting expectation and conditioning outcomes. In other words, whilst the conditioning procedure went towards a given direction, expectation was manipulated in the opposite direction. In this study, in one group of subjects a pharmacological pre-conditioning with ketorolac, a nonopioid analgesic, was performed for two days in a row and then ketorolac was replaced with a placebo on the third day along with verbal suggestions of analgesia. This procedure induced a robust placebo analgesic response. In order to see whether this placebo response was due to the pharmacological pre-conditioning, in a second group of subjects the same pre-conditioning procedure with ketorolac was carried out but the placebo was given on the third day along with verbal suggestions that the drug was a hyperalgesic agent. These verbal instructions were enough not only to block placebo analgesia completely, but also to produce hyperalgesia. These findings clearly show that expectation of hyperalgesia antagonized the pharmacological conditioning completely.

In the same study (Benedetti et al., 2003), Parkinson patients implanted for deep brain stimulation were tested for the velocity of movement of their right hand according to a double-blind experimental design in which neither the patient nor the experimenter knew whether the stimulator was turned off. The velocity of hand movement was assessed by means of a movement analyzer. The stimulator was turned off several times (at 4 and 2 weeks) before the test session. Each time the velocity of movement was measured just before the stimulator was turned off and 30 minutes later. On the day of the experimental session, the stimulator was maintained on but the patients were told that it had been turned off, so as to induce negative expectations of motor performance worsening (nocebo procedure). It was found that, although the stimulator was on, motor performance worsened and mimicked the worsening of the previous days. Importantly, this nocebo bradykinesia could be prevented completely by verbal suggestions of good motor performance (placebo procedure). Therefore, as occurs for pain, in this case also, motor performance can be modulated in two opposite directions on the basis of positive and negative expectations about motor performance.

Besides pain and motor performance, the same study (Benedetti et al., 2003) also tested hormone secretion, as already described above. Whereas verbally induced expectations of hormonal changes did not produce any effect on growth hormone and cortisol secretion, conditioning with sumatriptan induced placebo responses that mimicked the effects of sumatriptan itself, regardless of what the subjects expected (see the previous section). Therefore, in contrast to pain and Parkinson's disease, whereby expectation modulates the direction of the conditioning procedure, in hormonal secretion the direction of conditioning cannot be modulated by conscious expectations.

These data suggest that conscious expectation and unconscious conditioning may be at work in different conditions. In particular, conscious expectation of a future outcome seems to be important in the modulation of conscious processes, like pain and motor performance, whereas unconscious conditioning appears to be involved in unconscious physiological processes, such as hormone secretion. Therefore, the conditioned placebo hormonal responses indicate that unconscious conditioning is possible in humans, and cognitive components do not need to be always present.

Re-shaping neuronal activity through verbal suggestions of clinical improvement

Recent research aimed to identify possible cellular changes following placebo administration. In Parkinson's disease, whereby expectations and verbal suggestions of clinical benefit play a critical role in placebo responsiveness (see above), part of the neuronal circuit of the basal ganglia during a placebo response was investigated by means of microrecording techniques in awake patients who were undergoing electrode implantation for deep brain stimulation (Benedetti et al., 2004; Benedetti et al., 2009). In those patients who showed a clinical placebo response, there was a decrease in firing rate in subthalamic nucleus (STN) neurons, together with the disappearance of bursting activity, whereas placebo non-responders did not show these changes (Benedetti et al., 2004). By performing a detailed analysis of the output regions of STN, the same authors were able to identify part of the circuit that is involved in the anti-parkinsonian placebo response. In fact, the STN activity decrease was associated to a decrease in the substantia nigra pars reticulata (SNr) and an increase in the ventral anterior (VA) and anterior ventral lateral (VL_a) thalamus, thereby showing that placebo bridges STN and SNr activity whereas it increases VA-VL_a activity. By contrast, placebo non-responders showed either a lack of changes in this circuit or partial changes in the STN only. Thus, changes in activity in the whole basal ganglia-VA-VL_a circuit appear to be

important in order to observe a clinical placebo improvement (Benedetti et al., 2009). This circuit is likely to be a part of a more complex circuitry that is modified by a placebo treatment.

Therefore, verbal suggestions of benefit, which represent the basis of a placebo intervention, can reverse the malfunction of a complex neuronal circuit. This may have profound implications for psychotherapy, because the enhancement of expectations through verbal suggestions may indeed induce specific changes in the brain, thus placing psychotherapy into a therapeutic context which *per se* is capable of modifying the patient's brain.

Conclusions

Placebo responses represent the responses of the patient's brain to a variety of psychological factors and social stimuli. As a whole, these psychological/social factors constitute the ritual of the therapeutic act, through which the patient understands that he is being treated for his illness and, consequently, expects a clinical benefit. The therapeutic act may induce several changes in the patient's brain, including the release of specific neurotransmitters and the re-shaping of neuronal activity. These effects make the placebo response an excellent model to understand the intricate relationship between a complex mental activity, the nervous system and different physiological processes of the body. Furthermore, placebo responses represent a straightforward example of how verbal suggestions of benefit can induce changes at the cellular level, thereby placing psychotherapy into the neurobiological domain.

References

- Ader, R. and Cohen, N. (1975). Behaviorally conditioned immunosuppression. *Psychosomatic Medicine*, 37, 333–340.
- Amanzio, M. and Benedetti, F. (1999). Neuropharmacological dissection of placebo analgesia: Expectation activated opioid systems versus conditioning-activated specific subsystems. *Journal of Neuroscience*, 19, 484–494.
- Benedetti, F. (2008a). Mechanisms of placebo and placebo-related effects across diseases and treatments. *Annual Review of Pharmacology and Toxicology*, 48, 33–60.
- Benedetti, F. (2008b). *Placebo effects: Understanding the mechanisms in health and disease*. New York: Oxford University Press.
- Benedetti, F., Amanzio, M., Vighetti, S. and Asteggiano, G. (2006). The biochemical and neuroendocrine bases of the hyperalgesic nocebo effect. *Journal of Neuroscience*, 26, 12014–12022.

- Benedetti, F., Pollo, A., Lopiano, L., Lanotte, M., Vaghetti, S. and Rainero, I. (2003). Conscious expectation and unconscious conditioning in analgesic, motor, and hormonal placebo/nocebo responses. *Journal of Neuroscience*, 23, 4315–4323.
- Benedetti, F., Colloca, L., Torre, E., Lanotte, M., Melcarne, A., Pesare, M., Bergamasco, B. and Lopiano, L. (2004). Placebo-responsive Parkinson patients show decreased activity in single neurons of subthalamic nucleus. *Nature Neuroscience*, 7, 587–588.
- Benedetti, F., Lanotte, M., Colloca, L., Ducati, A., Zibetti, M. and Lopiano, L. (2009). Electrophysiological properties of thalamic, subthalamic and nigral neurons during the anti-parkinsonian placebo response. *Journal of Physiology*, 587, 3869–3883.
- Colloca, L. and Benedetti, F. (2005). Placebos and painkillers: Is mind as real as matter? *Nature Reviews Neuroscience*, 6, 545–552.
- Colloca, L. and Benedetti, F. (2009). Placebo analgesia induced by social observational learning. *Pain*, 144, 28–34.
- de la Fuente-Fernandez, R., Ruth, T. J., Sossi, V., Schulzer, M., Calne, D. C. and Stoessl, A. J. (2001). Expectation and dopamine release: Mechanism of the placebo effect in Parkinson's disease. *Science*, 293, 1164–1166.
- Enck, P., Benedetti, F. and Schedlowski, M. (2008). New insights into the placebo and nocebo responses. *Neuron*, 59, 195–206.
- Goebel, M. U., Trebst, A. E., Steiner, J., Xie, Y. F., Exton, M. S. et al. (2002). Behavioral conditioning of immunosuppression is possible in humans. *FASEB Journal*, 16, 1869–1873.
- Kirsch, I. (1999). *How expectancies shape experience*. Washington, DC: American Psychological Association.
- Mayberg, H. S., Silva, J. A., Brannan, S. K. et al. (2002). The functional neuroanatomy of the placebo effect. *American Journal of Psychiatry*, 159, 728–737.
- Petrovic, P., Kalso, E., Petersson, K. M. and Ingvar, M. (2002). Placebo and opioid analgesia – Imaging a shared neuronal network. *Science*, 295, 1737–1740.
- Price, D. D., Finniss, D. G. and Benedetti, F. (2008). A comprehensive review of the placebo effect: Recent advances and current thought. *Annual Review of Psychology*, 59, 565–590.
- Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, 43, 151–160.
- Scott, D. J., Stohler, C. S., Egnatuk, C. M., Wang, H., Koeppe, R. A. and Zubieta, J. K. (2007). Individual differences in reward responding explain placebo-induced expectations and effects. *Neuron*, 55, 325–336.
- Scott, D. J., Stohler, C. S., Egnatuk, C. M., Wang, H., Koeppe, R. A. and Zubieta, J. K. (2008). Placebo and nocebo effects are defined by opposite opioid and dopaminergic responses. *Archives of General Psychiatry*, 65, 220–231.
- Stockhorst, U., Gritzmann, E., Klopp, K. et al. (1999). Classical conditioning of insulin effects in healthy humans. *Psychosomatic Medicine*, 61, 424–435.
- Vase, L., Robinson, M. E., Verne, G. N. and Price, D. D. (2005). Increased placebo analgesia over time in irritable bowel syndrome (IBS) patients is associated with desire and expectation but not endogenous opioid mechanisms. *Pain*, 115, 338–347.
- Wager, T. D., Scott, D. J. and Zubieta, J. K. (2007). Placebo effects on human μ -opioid activity during pain. *Proceedings of the National Academy of Science USA*, 104, 11051–11061.

SECTION IV

Expanding boundaries



Shadow Puppet 4. Susan Aldworth (2007) etching and aquatint, 50 × 35 cm

The paradoxes of creativity

Ashish Ranpura and Mark F. Lythgoe

Creativity is the crowning jewel of human cognition, the ability that most uniquely characterizes our species. At the heart of subjective accounts of creativity are three deeply paradoxical features. First, creative acts are widely considered acts of self-expression, yet by many accounts the creative experience is selfless and only partially volitional. Second, creative activity is often hypnotically engrossing, while psychological theories about creativity emphasize loose, defocused thinking. Third, loose associations during creative acts might be profligate and degenerate, but creative products are often highly specific and optimal. Here we present two novel hypotheses which might explain the paradoxical nature of subjective accounts of creativity. First, we suggest that creativity requires a “quiet mind,” simultaneously focused and disinhibited. Second, we describe how this paradoxical activity in the cerebral cortex might support the Darwinian selection engine, a previously-proposed network-level mechanism for organizing and developing mental representations.

Socrates once asked how a man might become shorter than another without losing height (Frankfurt, 2004). It's a shallow paradox, but by the great philosopher's own admission, dwelling on it makes you dizzy with possibilities. You, the clever reader, are even now racking your brain in a flurry of possible interpretations about why a 2,500 year old parlor trick should present such a deep philosophical problem.

Of the many ways in which our thinking might be described as computation, this desire to push and play with ambiguity remains deeply and essentially human. Maybe all creative acts, from tricks with language to artistic expressions and scientific insights, prove to us that we are individual souls. They affirm that a man must be more than the sum of his constituent chemistry, that insight and inspiration are far removed from the mundane machinations of cells and synapses. Defining a biology of creativity isn't just an ambitious project; it's a proposition that profoundly tests our belief in the biological basis of mental life.

We don't hope to fully resolve this crisis of faith in these pages. Rather, we hope that by carefully examining converging lines of evidence and argument, we

can sketch out a plausible mechanism that might enable a humble lump of nerve tissue to produce an astonishing range of divergent, creative thought. Where previous authors have attempted to characterize creativity in terms of the range and quality of creative output, here we take a difference approach. We will first describe a novel characterization of subjective creative experience in the form of three deep paradoxes. The first of these is a paradox of ego: creativity seems to require a loosening of egocentric thinking, but creative products are often deeply representative of their creator (Limb and Braun, 2008). The second paradox is one of focus: the creative moment itself is hypnotically engrossing, while creative activity generally requires loose, defocussed thinking (Eysenck, 1995; Greely et al., 2008). And finally, the paradox of quality: why do degenerate, loose and profligate associations often result in extremely specific, high-quality solutions (Andreasen, 2006)?

We propose two novel hypotheses to account for these paradoxes of creativity. First, we suggest that creativity requires a “quiet mind,” simultaneously focused and disinhibited. The quiet mind hypothesis is a novel proposal, and we will present both introspective and objective evidence in its favor. Second, we will describe how the quiet mind can provide the ideal environment for a Darwinian selection engine, a previously-proposed network-level mechanism for organizing and developing activity patterns in the cerebral cortex (Calvin, 1996).

The paradoxes of creativity encapsulate something quirky, something essentially human, that defines creative thinking. In this essay, we will argue that these paradoxes are also necessary consequences of a biological mind. By directly tackling the most baffling elements of creativity, perhaps we can find some insight into how a very physical brain produces this most spiritual, unique, and somehow mundanely human of all experiences.

There is a certain feverish and often uncomfortable intensity that surrounds a creative act. The ego comes unhinged. Consciousness, as we normally experience it, shifts. At the height of our creative powers, our sense of self begins to dissolve, and we lose ourselves in the act of creation. Techniques to help foster creativity focus explicitly on subverting the censorious voice of the ego to allow a dormant, subconscious self to emerge (Edwards, 1999). But why is creativity so difficult to achieve in the course of our normal, conscious experience?

To have any hope of addressing that question, we'll limit our discussion of consciousness to the process that's engaged to produce and interpret an internal monologue – the inner narrator. That's what most non-specialists mean when they refer to consciousness, and roughly corresponds to Block's concept of “access consciousness” (Block et al., 1997). We don't intend our definition to be comprehensive and absolute, but it allows us to make a distinction between

those mental processes that we can observe ourselves doing (we'll call these conscious actions) and those processes that we're unaware of doing (subconscious actions).

That distinction lies at the heart of subjective descriptions of creativity. The bulk of what we normally consider a creative process is internal and almost without ego; this lack of conscious access to creativity leads the more spiritually inclined to consider the creative act an expression of the soul, or even of the divine. More simply put, the majority of the creative process happens subconsciously, in the absence of an internal monologue, and often in a trance-like state characterized as "flow" (Csikszentmihalyi, 1990). Flow is an egoless rapture, a state of automatic action with its own peculiar feeling of selfless consciousness. The flow state is a good illustration of the first paradox of creativity: creativity is both an egoless process and a profound expression of the self.

Maybe some aspects of this ego paradox can be clarified with our definition of consciousness itself. We are interested in consciousness as the voice of an internal narrator, an inner monologue through which we comment on our own mental life. It follows that fully conscious perceptions need to be unambiguous, so that they can be represented in language. The narrow focus of attention should encompass the whole of the conscious thought. Subconscious perceptions, on the other hand, are characteristically ambiguous – they represent incomplete and sometimes contradictory information. For the mind to operate on subconscious representations, the language of the inner narrator must be suppressed, ambiguity maintained. We suggest that when the mind is able to suppress the linguistic ego while still focusing on an abstract thought process, the result is the weird sense of dissociation that seems to accompany deeply creative moments.

This simple observation gives us our first hint about a brain mechanism for creativity: the creative brain must allow mental representations to remain partially-formed and ambiguous for a prolonged period of time (Gardner, 1984).

The ego paradox demonstrates the importance of prolonged ambiguity, but creativity is also characterized by an intense, bizarrely obsessive focus on the creative task itself. This tension between unfocused ambiguity and obsessive focus on the task is at the heart of the creative experience. It creates a mental state that can be fragile, easily lost, and even threatening. The great Russian composer Pyotr Tchaikovsky wrote that, "In the midst of this magic process it frequently happens that some external interruption wakes me from my somnambulistic state... Dreadful indeed are such interruptions. Sometimes they break the thread of inspiration for a considerable time, so that I have to seek it again – often in vain... If that condition of mind and soul which we call inspiration lasted long without intermission, no artist could survive it" (Andreasen, 2006).

Tchaikovsky describes an overwhelming obsession in the creative moment, one so powerful that it can't be maintained for long. But he also describes that state as "somnambulistic," magical, and dreamlike. Tchaikovsky's paradoxical observation can be translated into a prediction about brain function: in the creative state, the brain has low inhibition for ideas themselves, but high inhibition for mental task switching. The creative person is obsessive, and good evidence demonstrates that at the extremes there is a relationship between creativity, mania, and obsessional disorders (Andreasen, 1987). But even in its most benign manifestations, the paradox of focus is a hallmark of deeply creative thinking.

What can this tension between obsessive focus and loose association tell us about the neurobiological mechanisms of creativity? Simply that there is a need for both a driver and a vehicle in creative cognition. Strong task-orientation ensures that the maximum number of cortical resources are devoted to driving the creative project, and indeed in non-pathological situations, creative breakthroughs seem to occur only after periods of intense concentration. This concentration represents a drive. At the same time, loose associations and ambiguity allow idea generation, unconstrained by filtering mechanisms or goals. An idea generating machine can be a vehicle for creativity. A paradoxically focused brain is one which is driven to generate the maximum number of task-relevant mental representations, without requiring goal-oriented attention.

So far we've described creativity as a free-flowing mental state. Ideas blossom, interact, and change in the fertile landscape of an attentive but defocused and egoless mind. This notion of creativity as principally a generative process is at the heart of the Torrance Tests of Creative Thinking, an influential set of cognitive instruments that define creativity in the modern neuroscientific sense (Lissitz and Willhoft, 1985; Torrance, 1966).

The Torrance Tests do a good job of measuring a certain type of creativity, the ability to generate original, divergent solutions to well-defined problems. But practical artistic and scientific creativity is more elusive. In those situations problems are ill-defined and their solutions aren't immediately recognizable. The most creative thinkers aren't necessarily the ones who generate the most ideas or the strangest ones. Practical creativity is about insight, the ability to frame both the problem and the solution in terms that work. Creative solutions are ones that seem obviously right to everyone in hindsight, but were never obvious beforehand – they tend to elicit the "why didn't I think of that" response from onlookers.

In other words, the central feature of practical creativity is a judgement about quality. The Torrance Tests fail to capture that critical faculty of separating insight from nonsense. Dream states, for example, are low-quality, incoherent, and highly divergent. Sometimes dreams do inspire creative solutions, but the creative act

itself happens when the sense in a dream is distilled from the nonsense. Distilling requires expertise applied to a generative process; ultimately, distilling is driven by a desire to enact a solution, while the generative process is driven by a desire to attend to the creative problem. Finding real creative insights requires original, divergent thinking coupled with an ability to recognize high-quality solutions.

This brings us to the final paradoxical aspect of creativity, the paradox of quality. How do large numbers of loose, low-quality, mental representations sum to a single, high-quality solution? How does the brain flip from divergent creative thinking to actual, focused creating?

Some insight into this strange problem may come from an even stranger source: Darwinian evolution. Evolution by natural selection also proposes a system of change that begins with profligate, degenerate variants which, after passing a rigorous fitness test, end up defining a specific and optimal solution. In the natural world, a “variant” is simply a creature that’s slightly genetically different to either of its parents. This difference may or may not help the creature’s chances of survival and reproduction. If the genetic difference does help, then the creature is more likely to reproduce – and therefore the novel genetic contribution is more likely to spread. The proof is in the genetic pudding: creatures that appear extremely well-adapted to their environment became that way through massive random mutation coupled with specific selection.

Previous theoretical work has described how a Darwinian selection engine might be implemented in the neurons that comprise the cerebral cortex (Calvin, 1996). These neurons are densely interlinked, electrically active cells that represent information both by the set of physical connections they make and by the temporal frequency of their signals. Collectively, the neurons of the cerebral cortex represent a sort of tangled “jungle,” a cellular territory with resources for sustaining only a limited number of simultaneous activity patterns. The cortical jungle can be colonized by several activity patterns, which might compete through a selection engine. In analogy with natural selection, a cortical selection engine depends on a few key mechanisms: a way to generate an activity pattern, a way to copy and spread that pattern, a way to mutate or degrade the copies, and a way for the degenerate patterns to compete for cortical representation.

This kind of a Darwinian engine could be an excellent generator of creativity. With the drive of obsessive focus, the engine is pushed to build, mutate, and copy a large number of divergent representations of both problems and solutions. Each representation is a pattern of activity in the cerebral cortex, and the cortex is limited in the number of representations it can support simultaneously. So each representation – each creative possibility – competes for representational space. But in creative individuals, the representational space isn’t a blank slate. Biases in

the cortex, the residue of experience and training, favor certain activity patterns over others. Those activity patterns are copied more frequently, and represented more often in cortex. Solutions which are too far afield from experience and from each other soon wither and die out. Successful mutants help to create even more successful mutants, until eventually a single mental representation comes to dominate the jungle. We propose that the paradox of quality is resolved through an unsupervised, unconscious accumulation of degenerate changes, superimposed on a background of experience and knowledge.

There are times, however, when the background rule of expertise is turned on its head. Instances in which a 56-year-old housewife having no experience in art suddenly and compulsively begins to paint beautiful and sophisticated scenes (Miller et al., 1998). Or a 51-year-old builder, a practical man with no inclination towards esoteric expression, suddenly and obsessively begins to fill his house with sculpture and poetry (Lythgoe et al., 2005). In these rare and extraordinary stories, men and women have mostly suffered from damage to the frontal lobes from frontotemporal dementia (FTD), or a brain injury of some form.

FTD is an abnormal age-related change in brain structure, and initially it is associated with personality change and behavioral disinhibition. As the disease progresses it may affect a patient's ability to speak and to concentrate. And although most patients with brain injuries don't develop recognizably creative artistic tendencies, a few become profoundly artistically productive (Pollak et al., 2007). Some become accomplished visual artists (Lythgoe et al., 2005; Miller et al., 1998), others develop astonishing musical abilities (Miller et al., 2000), still others become writers and poets (Mendez, 2005). In all cases, the artistic productivity seems to reflect the paradoxes of ordinary creativity: it is egoless, focused, and, given its spontaneous nature, surprisingly high-quality. But in FTD patients, this creativity is often a strangely preserved island of ability in a rising sea of struggle and loss.

Why should patients with an overall deterioration in brain function suddenly exhibit an improved ability to think creatively? A clue to the answer might be found by examining the range of frontal lobe dysfunctions. Previous studies have demonstrated that patients with frontal lobe disturbances, including FTD, stroke, and traumatic brain injury, do worse at tasks that require egocentric thinking, like making judgements about self-relevant adjectives (Kelley et al., 2002). In addition, a long history of neuropsychiatry has shown that the frontal lobes are involved in task switching, and that frontal lobe damage often results in obsessional thinking and perseveration. So paradoxically, mild to moderate frontal lobe dysfunction may actually supercharge the Darwinian creativity engine by suppressing the ego and prolonging the longevity of degenerate mental representations, while at the same time promoting an intensely single-minded task-orientation. The resulting

state of mind was captured well by one of our patients, who reported an absolutely relentless flood of associations and ideas; as he put it, “Even when looking at my kitchen floor I see hundreds of different faces which I need to draw, paint or write about.” (Lythgoe et al., 2005).

The flood of associations that can result from frontal lobe injuries seems to be the result of a disinhibited mind (Pollack et al., 2007), what previous writers have called lowered latent inhibition (Carson et al., 2003). Latent inhibition may be described as the capacity of the brain to filter from conscious awareness stimuli previously perceived as irrelevant (Lubow and Moore, 1959). The latent inhibition theory fits well with anecdotal observations that creative individuals are often unusual or eccentric; they simply don’t self-censor to the same extent as others. The theory also does an excellent job of explaining the rapture of the creative experience, the obsessive focus and the relentless stream of ideas. Suppressing some frontal lobe functions, either through injury or through artificial techniques, seems to allow others to be unmasked (Snyder et al., 2003).

Like the Torrance Tests, the theory of latent inhibition emphasizes idea generation as the defining element of creative thought. What the theory leaves out is the element of insight. Why is it that creative acts are not simply bizarre, but in some cases that they are immediately recognizable as innovative and insightful? The productively creative individual has the ability to distill sense from nonsense. To see what everyone else has already seen in a new and useful way. In short, creative individuals are able to create.

The advantage of a Darwinian engine theory is that it makes the last explanatory leap, from creative thought to creative action. As the engine ticks over, generating ideas, mutating them, and allowing them to compete with one another, eventually a single solution comes to dominate. The dominance of that solution is influenced by previous experience, by learned parameters already present in the cortical environment. The productivity of the engine can be increased in a number of ways. It could happen through an obsessional drive, or by increasing the time that ideas have to compete for cortical dominance. It could happen by changing the rate of mutation, distorting cortical representations more or less. And finally, the creative output of the process can be increased through learning, observation and experience – those elements that ultimately switch the brain’s Darwinian engine from a drive to contemplate to an ability to create.

The three paradoxes of creativity, the paradox of ego, of focus, and of quality, are defining features of this quintessentially human experience. Evidence from neuroscience can now support the argument that frontal lobe functions may underlie all three paradoxes. Reduced frontal lobe functions might explain a suppression of egocentric thinking and the experience of flow; obsessive focus and

the desire to avoid distraction; and finally, the lifting of widespread inhibition to harness the power of a Darwinian selection engine. When these influences are tempered by expertise and insight, the ability to engage and the desire to act, the humble biology of our brains and bodies can produce a marvelous outpouring of soul and spirit.

References

- Andreasen, N. C. (1987). Creativity and mental illness: Prevalence rates in writers and their first-degree relatives. *American Journal of Psychiatry*, 144, 1288–92.
- Andreasen, N. (2006). *The creative brain: The science of genius*. New York: Penguin.
- Block, N. J., Flanagan, O. J. and Güzelde, G. (1997). *The nature of consciousness: Philosophical debates*. Cambridge, MA: MIT Press.
- Calvin, W. H. (1996). *The cerebral code: Thinking a thought in the mosaics of the mind*. Cambridge, MA: MIT Press.
- Carson, S. H., Peterson, J. B. and Higgins, D. M. (2003). Decreased latent inhibition is associated with increased creative achievement in high-functioning individuals. *Journal of Personality and Social Psychology*, 85, 499–506.
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper and Row.
- Edwards, B. (1999). *The new drawing on the right side of the brain*. New York: Jeremy P. Tarcher/Putnam.
- Eysenck, H. J. (1995). *Genius: The natural history of creativity*. Cambridge: Cambridge University Press.
- Frankfurt, H. G. (2004). *The reasons of love*. Princeton, N.J.: Princeton University Press.
- Gardner, H. (1984). *Art, mind, and brain: A cognitive approach to creativity*. New York: Basic books.
- Greely, H., Sahakian, B., Harris, J., Kessler, R. C., Gazzaniga, M., Campbell, P. and Farah, M. J. (2008). Towards responsible use of cognitive-enhancing drugs by the healthy. *Nature*, 456, 702–705.
- Kelley, W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S. and Heatherton, T. F. (2002). Finding the self? An event-related fmri study. *Journal of Cognitive Neuroscience*, 14, 785–94.
- Limb, C. J. and Braun, A. R. (2008). Neural substrates of spontaneous musical performance: An FMRI study of jazz improvisation. *Plos ONE*, 3, e1679.
- Lissitz, R. W. and Willhoft, J. L. (1985). A methodological study of the torrance tests of creativity. *Journal of Educational Measurement*, 22, 1–111.
- Lythgoe, M. F., Pollak, T. A., Kalmus, M., de Haan, M. and Chong, W. K. (2005). Obsessive, prolific artistic output following subarachnoid hemorrhage. *Neurology*, 64, 397–398.
- Mendez, M. F. (2005). Hypergraphia for poetry in an epileptic patient. *Journal of Neuropsychiatry and Clinical Neuroscience*, 17, 560–561.
- Miller, B. L., Boone, K., Cummings, J. L., Read, S. L. and Mishkin, F. (2000). Functional correlates of musical and visual ability in frontotemporal dementia. *British Journal of Psychiatry*, 176, 458–463.

- Miller, B. L., Cummings, J., Mishkin, F., Boone, K., Prince, F., Ponton, M. and Cotman, C. (1998). Emergence of artistic talent in frontotemporal dementia. *Neurology*, 51, 978–982.
- Pollak, T., Mulvenna, C. and Lythgoe, M. (2007). De novo artistic behaviour following brain injury. *Frontiers of Neurology and Neuroscience*, 22(R), 75.
- Snyder, A. W., Mulcahy, E., Taylor, J. L., Mitchell, D. J., Sachdev, P. and Gandeia, S. C. (2003). Savant-Like skills exposed in normal people by suppressing the left fronto-temporal lobe. *Journal of Integrative Neuroscience*, 2, 149–158.
- Torrance, E. P. (1966). *The torrance tests of creative thinking*. Princeton, N.J.: Personal Press.

Potential contributions of research on meditation to the neuroscience of consciousness

Antoine Lutz

Exploring the initial findings of neuroscientific research on meditation, new horizons of further inquiry in consciousness research are apparent. While such studies of contemplative practices are still in their infancy, early findings promise to contribute in three key areas. These include: Neuroplasticity – physiological and psychological indices of short and long terms responses of the brain circuits that underlie complex mental functions associated with specific types of meditation techniques; Mind body Interactions – revealing mechanisms by which such training may exert beneficial effects on physical health; and Subjectivity – well developed introspective skills of practitioners potentially shedding new light on the neural counterpart of subjectivity.

How does one define meditation? How does contemplative neuroscience contribute to the study of consciousness? These questions are not easily answered, in part because the term “meditation” encompasses many activities and in part because of the paucity of empirical data currently available in this field. For these reasons, this chapter remains largely programmatic (for details see Lutz et al., 2006; Lutz et al., 2008). Its overall goal is to explore the initial findings of neuroscientific research on meditation; in doing so, it also suggests new horizons of further inquiry in conscious research. This chapter has four sections. The first section clarifies meditation as an explanandum, provides a brief overview of some standard meditation practices and proposes three methodological advantages to collaborate with practitioners to study consciousness. These methodological advantages appear particularly relevant in relation to three neuroscientific agendas: neuroplasticity, the interaction of mind and body, and the possibility of neural counterparts to subjective experience. These three agendas are discussed in the remaining three sections.

Meditation as an explanandum

The term ‘meditation’ refers to a broad variety of practices, ranging from techniques designed to promote relaxation to exercises performed with a more far-reaching goal such as a heightened sense of well-being. It is thus essential to be specific about the type of meditation practice under investigation. Failure to make such distinctions would be akin to the use of the word “sport” to refer to all sports as if they were essentially the same. In order to narrow the explanandum to a more tractable scope, this article uses Buddhist contemplative traditions as a paradigmatic framework. The reason, in brief, is that unlike many contemplative traditions, Buddhist traditions tend to offer extensive, precisely descriptive and highly detailed theories about their practices in a manner that lends itself readily to appropriation into a neuroscientific context. Despite the variety of contemplative practices encompassed by the term “meditation” within Buddhism, most Buddhist traditions use a term for meditation that correlates with the Sanskrit term *bhāvanā*, literally, “causing to become.” In Tibetan traditions, the usual translation for *bhāvanā* is *gōm* (*sgom*), which roughly means “to become habituated to,” or “to become familiar with.” These Buddhist traditions share two axioms articulated in highly detailed scholastic texts such as the *Abhidharma*: (1) that a central goal of Buddhist practice is the elimination of suffering, where “suffering” coarsely refers to the negative affects associated with physical pain, aging, neurosis or social stress; and (2) that any effective method to eliminate suffering must involve changes in one’s cognitive and emotional states, in particular one’s self-centered habits (Gethin, 1998). Meditation is thus conceptualized here as a family of complex emotional, conceptual and attentional regulatory training regimes developed for various ends, including the cultivation of well-being and emotional balance.

The most commonly practiced style of meditation consists of focusing or stabilizing concentration on an object such as the breath (this concentrative part is here labeled Focus Attention (FA) meditation. See Lutz et al., 2008 for details). By cultivating a form of meta-attention, or non-reactive awareness, this technique further enables the practitioner to be more anchored in one’s experience from moment to moment and to gain insight into one’s habits and assumptions about identity and emotions (this later stage is here labeled Open Monitoring (OM) meditation, see Lutz et al., 2008). This style of meditation is often called Vipassana meditation (*Vipaśyanā* in Sanskrit, or *Vipassanā* in Pāli) or mindfulness meditation. It is traditionally paired with practices such as compassion and loving-kindness meditations. These meditations often combine an emotional component, such as the invocation of a state of compassion, with the development of cognitive schemas, such as the cultivation of a sense of equanimity toward others. The

purpose of these compassion-based meditations is to cultivate altruistic behaviors and positive effects and to counteract self-centered tendencies. The Buddhist soteriological project is thus intrinsically embodied in an explicit effort to directly examine, understand and transform one's mind through meditative training.

The neuroscientific study of advanced practitioners of the aforementioned styles of meditation may contribute to the study of consciousness through, among others, three methodological hypotheses: (1) Advanced practitioners can generate new data that would not exist without sustained mental training. These data encompass either meditative states or traits induced by meditation. (2) Advanced practitioners can robustly reproduce specific features in experience as cultivated in a given meditative practice. This reproducibility makes those features scientifically tractable. (3) Advanced practitioners provide more refined first-person descriptions of their experiences than naïve subjects. Thus, the neurophysiological counterpart of these first-person accounts can be defined, identified and interpreted more easily by the experimentalist. These hypotheses appear particularly relevant in relation to three neuroscientific agendas: neuroplasticity, the interaction of mind and body and the possibility of neural counterparts to subjective experience.

Transforming the mind and neuroplasticity

In the Buddhist tradition, emotions, attention and introspection are viewed as ongoing and labile processes to be understood and studied as skills that can be trained, similar to other human domains like music, mathematics or sports. As a result, the methods employed by Buddhist contemplative practices resonate with widely accepted developmental models of basic cognitive processes. This basic stance also reflects another well-documented theory: that experience changes the brain, a process known as neuroplasticity. The current findings on neuroplasticity raise the possibility that trainings that are specifically designed to cultivate these skills will produce specific alterations in brain function and structure. Several recent studies have reported expertise-related changes in attentional processing and brain structures in those proficient in attention-based meditation. In one study (Lutz et al., 2004), we showed that, in comparison to controls, the ratio of fast-frequency electroencephalogram (EEG) activity (25–42 Hz) to slow oscillatory activity (4–13 Hz) is initially higher on average across electrodes in the resting baseline before meditation for advanced practitioners. During objectless compassion meditation, this difference increases sharply over medial fronto-parietal electrodes and remains higher than the initial baseline in the post-meditative baseline.

These changes across group and time support the view that meditative practice induces short-term and possibly long-term functional changes in brain oscillatory processes. Similarly, Carter et al. (2005) found that Tibetan Buddhist monks were able to perceive a stable, superimposed percept of two dissimilar, competing images presented to separate eyes for a longer duration both during and after a Focused Attention (FA) meditation but not a form of 'compassion' meditation. These behavioral findings obtained using a 'binocular rivalry' task support the hypothesis that extensive training in FA meditation may improve the practitioner's ability to sustain attention on a particular object for a prolonged period of time. Building on this finding, we recently investigated the impact of Vipassana meditation on the moment-to-moment stability of attention by quantifying the trial-to-trial variability of both reaction time and consistency of brain responses, as measured by electroencephalography (EEG) during a selective attention task (Lutz et al., 2009). This measure of variability in both the behavioral and brain response is particularly relevant to measure ongoing fluctuations in attention stability, which are thought to reflect competitive interactions between task-related and task-unrelated processes, such as mind-wandering. For instance, Attention Deficit Hyperactivity Disorder (ADHD) patients typically show larger intra-individual variability in reaction time than controls during performance of sustained attention tasks. We showed that three months of rigorous training in Vipassana meditation improved people's ability to stabilize attention on target tones, as indicated by both a measure of consistency of brain response and reduced reaction time variability. Furthermore, those individuals who showed the greatest increase in neural response consistency showed the largest decrease in reaction time variability.

In the same study we further investigated the impact of OM meditation on mental and brain functions. In Vipassana, OM meditation is typically practiced once the focus of attention is more stabilized. OM meditation fosters nonreactive awareness of the stream of experience without deliberate selection of a primary object. Intensive practice can be expected to reduce the elaborative thinking that would be stimulated by evaluating or interpreting a selected object. In line with this prediction, Slagter et al. (2007) further found, during an attentional blink task, reduced elaborative processing of the first of two target stimuli (T1 and T2) presented in a rapid stream of distracters, as indicated by a smaller T1-elicited P3b, a brain potential index of resource allocation (Slagter et al., 2007). Moreover, this reduction in resource allocation to T1 was associated with improved detection of T2. Because participants were not engaged in formal meditation during task performance, these results provide support for the idea that one effect of an intensive training in OM meditation might show reduction in the propensity to 'get stuck' on a target, as reflected in less elaborate stimulus processing and the

development of efficient mechanisms to engage and then disengage from target stimuli in response to task demands.

These changes across group and time support the view that meditative practice induces short-term, and possibly long-term, functional changes in brain oscillatory processes. Consistent with these findings, several groups have started reporting meditation-related structural changes using magnetic resonance imaging. In one study, cortical brain regions associated with attention, interoceptive and sensory processing were found to be thicker for a group of mid-range practitioners than for matched controls (Lazar et al., 2006). In another study, Pagnoni et al. found age effects on gray matter volume and attentional performance among Zen meditation practitioners (Pagnoni et al. 2007). While control subjects displayed the expected negative correlation of both gray matter volume and attentional performance with age, meditators did not show a significant correlation of either measure with age.

Interaction mind and body

Although there are many popular claims about the health benefits of meditation and contemplative practice, there is relatively little that is reliably known about this potentially crucial issue. It is established that there is bi-directional communication between the brain and peripheral biological systems, and that this communication proceeds along three basic routes: the autonomic nervous system, the endocrine system and the immune system. Some conditions of the peripheral biological systems may be affected by meditative practices because those conditions – such as an illness – are susceptible to modulation by the autonomic, endocrine and/or immune pathways involved in brain-periphery communication. Davidson et al. examined this possibility on individuals randomly assigned to a secularized, 8-week course in meditation called the Mindfulness Based Stress Reduction (MBSR) program (Kabat-Zinn, 1985) or a wait-list control group. This program is routinely used in a clinical environment as a complementary treatment to a variety of illnesses such as depression, chronic pain or anxiety (Kabat-Zinn, 1985). After this training, meditators exhibited a significantly greater antibody response to the influenza vaccine than controls. The individual differences in patterns over central electrode locations, interpreted as reflecting dispositional positive affect (Davidson et al., 2003), were also positively associated with differences in antibody response to the influenza vaccine (Davidson et al., 2003). The long-term consequence of this training is still unknown and is being explored. Carlson et al. (2007), for instance, studied the ongoing effects of MBSR on biomarkers of stress in a population of cancer patients. 59, 51, 47 and 41 patients were assessed

pre- and post-intervention and at 6- and 12-month follow ups. The participation in this intervention was associated with enhanced quality of life, decreased stress symptoms and altered cortisol and immune patterns consistent with less stress and mood disturbance. The long-term question is to evaluate whether repeated practice in such techniques could result in enduring changes in the functions and structures of peripheral biology and whether these changes are specific to selected practices (for a preliminary study on the effect of compassion meditation on the immune response see Pace et al., 2009).

Neural counterpart of subjectivity

As stated above, meditation training may develop introspective skill. Thus, the neurophysiological counterpart of the first-person accounts from advanced practitioners may be identified more easily by the experimentalist. Furthermore, various meditative practices induce a wide variety of altered states of consciousness. Some forms of Buddhist objectless meditation, for example, attempt to de-emphasize the explicit sense of narrative or autobiographical self, that encompass categorical or moral judgment, emotions, anticipation of the future and recollections of the past, so as to further enhance access to a minimal subjective sense of “I-ness” in experience, or ipseity (for details see Lutz et al., 2006). The collaboration with long-term Buddhist practitioners is thus of great potential interest in the study of the physical correlates of subjectivity or the self.

As OM meditation involves the cultivation of awareness of the subjective features of a given moment, such as its emotional tone it is conceivable that it engages processes involved in interoception, or the perception of internal bodily responses (Lutz et al., 2008). These processes rely on meta-representations in the brain of homeostatic afferent activity (e.g. temperature change or pain), in particular in the anterior insula, somatosensory cortex and anterior cingulate cortices (Craig, 2007). To test this idea behaviorally, we compared two groups of meditators (Tibetan Buddhist and Kundalini) to an age- and body mass index-matched group of nonmeditators (Khalsa et al., 2008). Contrary to our prediction, we found no evidence that meditators were superior to nonmeditators in the heartbeat detection task, across several sessions and respiratory modulation conditions. Contrary to this pilot study and in line with the idea that OM meditation engages processes involved in monitoring one’s body state, a recent study found greater activity in this neural circuitry during a monitoring state relative to a narrative generation state in participants who attended an eight-week course incorporating OM meditation (MBSR) compared with a group of controls (Farb et al., 2007). This promising avenue of research critically needs further investigation.

Finally, while this discussion has focused largely on long-term Buddhist practitioners, it is crucial to note the importance of future research on novices or on practices from other contemplative traditions (for an example of a neuroimaging study on Christian contemplatives, see Beauregard et al., 2008). Overall the neuroscientific study of contemplative practices is still in its infancy, but the early findings promise (1) to underscore the plasticity of the brain circuits that underlie complex mental functions; (2) to reveal the mechanisms by which such training may exert its beneficial effects on physical health and (3) to potentially shed new light on the neural counterpart of subjectivity.

References

- Beauregard, M., and Paquette, V. (2008). EEG activity in carmelite nuns during a mystical experience. *Neuroscience Letters*, 444, 1–4.
- Carlson, L. E., Speca, M., Faris, P. and Patel, K. D. (2007). One year pre-post intervention follow-up of psychological, immune, endocrine and blood pressure outcomes of mindfulness-based stress reduction (MBSR) in breast and prostate cancer outpatients. *Brain, Behavior, and Immunity*, 21, 1038–1049.
- Carter, O. L., Presti, D. E., Callistemon, C., Ungerer, Y., Liu, G. B. and Pettigrew, J. D. (2005). Meditation alters perceptual rivalry in tibetan buddhist monks. *Current Biology*, 15, R412–3.
- Craig, A. D. (2007). Interoception and emotion: A neuroanatomical perspective. In M. Lewis, J. M. Haviland-Jones & L. F. Barrett (Eds.), *In handbook of emotion* (3rd edition). New-York, NY: Guilford Press.
- Davidson, R. J., Kabat-Zinn, J., Schumacher, J., Rosenkranz, M., Muller, D., Santorelli, S. F., Urbanowski, F., Harrington, A., Bonus, K. and Sheridan, J. F. (2003). Alterations in brain and immune function produced by mindfulness meditation. *Psychosomatic Medicine*, 65, 564–570.
- Farb, N. A., Segal, Z. V., Mayberg, H., Bean, J., McKeon, D., Fatima, Z. et al. (2007). Attending to the present: Mindfulness meditation reveals distinct neural modes of self-reference. *Social Cognitive and Affective Neuroscience*, 2, 313–322.
- Gethin, R. (1998). *The foundations of Buddhism*. Oxford: Oxford University Press.
- Kabat-Zinn, J., Lipworth, L. and Burney, R. (1985). The clinical use of mindfulness meditation for the self-regulation of chronic pain. *Journal of Behavioural Medicine*, 8, 163–190.
- Khalsa, S. S., Rudrauf, D., Damasio, A. R., Davidson, R. J., Lutz, A. and Tranel, D. (2008). Interoceptive awareness in experienced meditators. *Psychophysiology*, 45, 671–677.
- Lazar, S. W., Kerr, C. E., Wasserman, R. H., Gray, J. R., Greve, D. N., Treadway, M. T., McGarvey, M., Quinn, B. T., Dusek, J. A., Benson, H., Rauch, S. L., Moore, C. I. and Fischl, B. (2005). Meditation experience is associated with increased cortical thickness. *Neuroreport*, 16, 1893–1897.
- Lutz, A., Greischar, L. L., Rawlings, N. B., Ricard, M. and Davidson, R. J. (2004). Long-term meditators self-induce high-amplitude gamma synchrony during mental practice. *Proceedings of the National Academy of Science USA*, 101, 16369–16373.

- Lutz, A., Dunne, J. and Davidson, J. R. (2006). Meditation and the Neuroscience of Consciousness: An Introduction. In M. M. a. T. E. Zelazo P. (Ed.), *The Cambridge Handbook of Consciousness*: Cambridge University Press.
- Lutz, A., Slagter, H. A., Dunne, J. D. and Davidson, R. J. (2008). Attention regulation and monitoring in meditation. *Trends in Cognitive Sciences*, 12, 163–169.
- Lutz, A., Slagter, H. A., Rawlings, N. B., Francis, A. D., Greischar, L. L. and Davidson, R. J. (2009). Mental training enhances attentional stability: Neural and behavioral evidence. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 29, 13418–13427.
- Pace, T. W., Negi, L. T., Adame, D. D., Cole, S. P., Sivilli, T. I., Brown, T. D. et al. (2009). Effect of compassion meditation on neuroendocrine, innate immune and behavioral responses to psychosocial stress. *Psychoneuroendocrinology*, 34, 87–98.
- Pagnoni, G. and Cekic, M. (2007). Age effects on gray matter volume and attentional performance in zen meditation. *Neurobiology of Aging*, 28, 1623–1627.
- Slagter, H. A., Lutz, A., Greischar, L. L., Francis, A. D., Nieuwenhuis, S., Davis, J. M. et al. (2007). Mental training affects distribution of limited brain resources. *PLoS Biology*, 5(6), e138.

Self-induced altered states of consciousness

Bangalore N. Gangadhar and Naren P. Rao

Consciousness can be altered willfully. There are different methods to induce altered state of consciousness including psychological means. While some such induced states are pathological many others are not. Indian and other eastern concepts emphasize one can induce “higher states of consciousness” oneself. Physiological changes are observed in these altered states of consciousness that point to a neural basis.

Consciousness has attracted attention of scientists from various disciplines and has been called the last frontier of science. However, there is no uniform definition of consciousness. It is a multi-faceted concept with different definitions but has been used mainly in three meanings: consciousness as awareness, consciousness as experience and consciousness as mind.

From a medical viewpoint, in neurology, consciousness is equated with waking state, with the abilities to perceive, interact, and communicate with others in the integrated manner. This ranges from waking through sleep into coma. However, in addition to its outward characteristics, consciousness has inner characters too, which are qualitative. This inner experience is beyond the normal awareness and is explained in Eastern philosophical traditions as “being aware of one’s awareness”. Neurophysiologically, consciousness has two components, namely vigilance (the level of wakefulness) and awareness. Awareness is further divided into conscious perception of the environment and of our inner world, that is self-awareness (Laureys, 2005).

Altered states of consciousness

As there is no uniform definition of consciousness, the term ‘altered states of consciousness (ASC)’ though encountered frequently, has defied systematic definition and boundary. Various conditions, ranging from reversible to irreversible have been described. Though some of the pathological conditions are irreversible, most

of the authors restrict the term ASC to reversible, short term conditions (Farthing, 1992). The Altered states of consciousness consortium, has focused mainly on clinical conditions or those states induced by physiological and psychological methods. ASC can be classified based on whether it is spontaneous or induced. Spontaneous altered states of consciousness occur in health during daily activities and include states of drowsiness, hypnagogic states, sleep and daydreaming.

The induced variety can be autogenous (that is induced by self) or allogeous (that is induced by others). The mode of induction may vary namely, physical, physiological, pharmacological or psychological. Examples of Physically and Physiologically induced altered states of consciousness include respiratory maneuvers, sexual activity and orgasm, extreme environmental conditions, starvation, sensory deprivation, homogenization (a homogeneous perceptual field usually visual but may also be auditory) or overload. Those induced by Psychological means include myriads of conditions namely relaxation, different kinds of meditation (Pranayama, Sudarshan Kriya, Transcendental Meditation, Om meditation), hypnosis, biofeedback and rhythm-induced trance. Pharmacologically induced ASC are the result of use of psychoactive substance like LSD. From a medical point, ASC can be physiological or pathological. Physiologically induced states are as described above. Pathological ASC may be organic or functional. Epilepsy is an example for organic ASC and dissociation is an example for functional ASC.

In this review, we focus on self-induced altered states consciousness by psychological methods and briefly discuss about the Indian and eastern concepts of consciousness with a focus on 'higher states of consciousness'.

Phenomenologically, altered states of consciousness are heterogeneous. As seen above, the causes of ASC are multitude, so also the features of ASC. A number of dimensions have been proposed by different authors to define the characteristics of altered states of consciousness for example Farthing (Farthing, 1992) proposed a 14-dimension concept.

A four dimensional descriptive system has been proposed recently to describe the phenomenology and physiologic mechanisms behind altered states of consciousness. Alterations in one or many of these can result in altered state of consciousness (Vaitl et al., 2005). The four dimensions are activation (readiness of an organism to interact with its physical or social environment), awareness span (variability of contents available for conscious processing), self-awareness, and sensory dynamics (sensory and perceptual component of subjective experience). Different methods of induction cause alterations in different dimensions in both the directions; the experiences are either enhanced or reduced. For example, sensory deprivation results in decrease in sensory dynamics and activation, but awareness span is increased. In meditation, activation and self-awareness is

reduced and awareness span and sensory dynamics varies according to the kind of meditation technique – it may be narrowed in concentrative meditation but widened in mindfulness meditation. In dissociation, while there is decrease in self-awareness other three dimensions are preserved. Among the four dimensions, activation is the basic and is nearest to the physiological description as it is related to changes in autonomic nervous system. However, the other dimensions are related to central and regulatory functions of higher brain regions.

These mechanisms are related to the clinical features of ASC as exemplified in two common forms of ASC, one being pathological – dissociation and the other being psychological self induced – meditation.

Dissociation. The essential feature of dissociative disorders is a disruption in the usually integrated functions of consciousness, memory, identity or perception of environment. Significant stressful event precedes the onset of symptoms. Different subtypes of dissociative disorders have been described; conversion disorder, dissociative amnesia, dissociative fugue and dissociative identity disorder. In dissociative stupor there is significant reduction or absence of voluntary movement along with disturbance of consciousness in which patient is neither asleep nor conscious. In dissociative amnesia, there is a loss of knowledge of personal identity with preservation of other information. In dissociative fugue, in addition to amnesia, an impulse to wander is present; the patient may travel away from home and at times to unusually long distances. Sometimes, a partial new identity may be assumed. The most interesting and widely publicized form of dissociation is dissociative identity disorder (earlier called as multiple personality disorder). In this there are two or more distinct identities, each with its own relatively enduring pattern of perceiving, relating to and thinking about the environment and self. Most of them lack memory of other personality but some know each other (Merskey and Piper, 1998). Thus, importantly, there is decreased self-awareness along with reduced activation and awareness span.

Meditation. There are different methods of yogic practice and each has different dimensions of altered state. Among the meditation techniques, Transcendental Meditation (TM) and Sudarshan Kriya Yoga (SKY) are widely examined. SKY consists of specific sequence of breathing patterns (pranayam) separated by brief periods of normal breathing: *Ujjayi* is the phase of slow strained breathing, *bhastrika* is forceful exhalation and a meditation phase of chanting *Om* (Janakiramaiah et al., 2000). This cyclical breathing is practiced with a rate of 60 cycles per min up to a period of 20 minutes. The procedures are synchronized with chanting of mantra “*so hum*”. The subjective experience is of excitation, followed by emotional calming with mental activation and alertness.

Neural basis of consciousness

Pioneering work of Moruzzi and Magoun revealed that ascending reticular activating system in brainstem and thalamus regulates the state of consciousness (Moruzzi and Magoun, 1949). However, later studies in animals showed that consciousness does not depend upon one structure: rather, multiple structures are involved in maintenance and modulation of wakefulness. These structures include cholinergic nuclei in brainstem and basal forebrain; locus coeruleus, the noradrenergic nuclei; histaminergic projection from posterior hypothalamus and dopaminergic and serotonergic pathways from brainstem. Functional imaging experiments in humans have also confirmed the involvement of these areas in maintenance of consciousness: during slow wave sleep there is decline in blood flow in rostral brainstem, thalamus, prefrontal cortex and cingulate cortex (Hofle et al., 1997).

Electrophysiological studies have shown that in wakeful state, thalamocortical neurons are tonically depolarized by cholinergic, noradrenergic and histaminergic inputs from the brainstem and hypothalamus, which block a hyperpolarizing potassium conductance. During sleep, there will be a simultaneous disinhibition of reticular nucleus of thalamus and reduction of inhibitory cholinergic input from brainstem resulting in a synchronized GABAergic thalamocortical inhibition. Interestingly, recent functional neuroimaging studies have suggested possible role of precuneus in self-consciousness and sense of agency. Precuneus and surrounding area have highest resting metabolic rates in non self-referential goal directed activities. Also, consistent with the above, there is hypometabolism in this area in altered states of consciousness like sleep, vegetative states and anesthesia (Cavanna and Trimble, 2006).

Neural basis of altered states of consciousness

Meditation. Transcendental meditation is associated with decreased heart rate and higher skin resistance. The changes may reflect the change in the state of consciousness of the subject from awareness to altered transcendental consciousness. Different changes like increase in alpha activity and coherence have been noted in EEG, during meditation (Fenwick et al., 1977). Some have proposed this as reflective of ordering of the mind but considered nonspecific by others.

Individuals who practiced meditation for a long term differed from those who practiced for only a short term with respect to theta activity over frontal region. Interestingly, there was also increased theta synchronization in left prefrontal

cortex and posterior association cortex suggesting an emotionally positive state. SKY results in increased brain perfusion, increased vigilance and slowing of heart rate due to parasympathetic dominance and people feel calmed. This subsequently, results in thalamic stimulation in turn quieting of frontal cortex. This is reflected in EEG as increased beta activity and increased alpha activity (Brown and Gerbarg, 2005).

Functional neuroimaging studies

With the advent of functional neuroimaging, a new window was opened to investigate the neurobiology of meditation and the dynamic neurophysiology of altered states of consciousness. Different techniques like positron emission tomography (PET), single photon emission computed tomography (SPECT) were used (Newberg and Iversen, 2003). Different forms of meditation have been examined namely yoga, tantric meditation, Tibetan meditation. Interestingly, though the techniques differed, most of them had activation of prefrontal cortex and hippocampus but in one study, there was decreased activation of parietal cortex (Newberg and Iversen, 2003). Recently Lazar and colleagues used functional magnetic resonance imaging (fMRI) in a study. They examined fMRI changes associated with kundalini meditation in five practitioners. There was activation of neural structures involved in attention (frontal and parietal cortex) and arousal/autonomic control (pregenual anterior cingulate, amygdala, midbrain and hypothalamus) on meditation. There was also a significant activation identified in the putamen, precentral and postcentral gyri as well as hippocampus/parahippocampus suggesting their role in meditative states (Lazar et al., 2005).

OM meditation. Meditation involves chanting of certain mantras while focusing on breathing. One of the important chanting used in meditation is 'OM'. Thus, OM meditation has attracted several researchers. In one study, the auditory middle latency evoked response potential obtained during OM meditation had shorter latency and larger amplitude suggestive of a possible midbrain-thalamic activation. There was also a lowering of sympathetic tone or increase in parasympathetic tone, which could be because of the effects of OM chanting on limbic circuits (Telles et al., 1994). Recently, we studied neurohemodynamic correlates of OM chanting using fMRI. Significant deactivation was observed in bilateral orbito-frontal, anterior cingulate, parahippocampal gyri, thalami, hippocampi and right amygdala. Together, these indicate limbic deactivation that may help quieting the emotions (Kalyani et al., 2009).

Self induced hypnosis. Hypnosis is both self-induced and induced by others. Hypnosis is reported to be associated with left dorsolateral prefrontal inhibition; neuropsychological studies have shown impaired letter fluency. EEG studies have reported higher delta and theta EEG activity posteriorly and alpha and beta-1 activity anteriorly during hypnosis, suggesting the co-occurrence of electrophysiological characteristics of lowered vigilance and increased attention (Katayama et al., 2007).

Relaxation. Relaxation predominantly has three components- focusing, passivity and receptivity. Focusing is the ability to concentrate on a single stimulus; passivity is the ability to refrain oneself from an active analytical thinking; receptivity is the ability to tolerate different contrasting experiences. Relaxation is associated with physiological changes such as decrease in sympathetic activity, resulting in bradycardia, decreased skin conductance, peripheral vasodilatation, decreased oxygen consumption (Shapiro and Lehrer, 1980). Imaging studies have shown activation of left superior temporal gyrus and anterior cingulate cortex (Critchley et al., 2001). Particularly, there was a strong relation between enhanced anterior cingulate activity and voluntary control of bodily arousal. Thus, relaxation, capable of reducing autonomic and central arousal can also be utilized to induce ASC.

Dissociation. Three neurobiological theories of dissociation have been proposed; neuro-chemical, epileptic model and hemispheric laterality model. Neurochemical theories indicate that stress increases release of glutamate and results in development of dissociative symptoms (Loewenstein and Putnam, 2005). This is based on the observation that NMDA receptor antagonists like ketamine, can result in release of glutamate and subsequently dissociation like symptoms. This is also supported by hippocampal volume loss in structural MRI studies. However, in addition to glutamate, serotonin and noradrenaline are also implicated as LSD and Yohimbine can also result in dissociative symptoms. The epileptic model is based on the observation of dissociative symptoms in patients with epilepsy during preictal and ictal stages and increased prevalence of temporal slow waves in dissociative disorder.

Thus, while some of the allogeous altered states may follow 'bottom-up' method with main effect on activation; autogenous altered states of consciousness probably follow a 'top-down' inhibition. That is, through different mechanisms, there is either exertion of self-control (as in meditation, relaxation) or loss of control (as in dissociation), which in turn results in changes in autonomic and cortical functions. This subsequently results in increase or decrease in one or more of the four domains of consciousness leading to altered state of consciousness (Vaitl et al., 2005).

Indian and other Eastern thoughts

Indian thought conceptualizes consciousness as a multilayered entity. It differentiates awareness from consciousness; basic level of consciousness is called as *chethana* or *pragna* (awareness). On this, the *chitta* or *manas* (mind) is created by the sensory and motor systems (*Indriyas*). *Manas* responds to the input from *indriyas* and decides whether to reject or accept these inputs. Above the *manas* lies the *buddhi* (intelligence) which discriminates and imposes decision on *manas*. Above the level of awareness comes 'self-awareness' or 'I'ness, which is being aware of oneself. The 'I' can perceive the action of mind and can restrain the mind.

According to Upanishads, the ancient Hindu scripts, consciousness evolves in each individual from physical plane to vital plane to mental plane. If one could rise above the mental plane they can reach the plane of creative intelligence (*budhi*) and then that of bliss (*ananda*). This is described in *Taittiriya Upanishad* as different layers of consciousness: physical (*anna-mayakosa*), psychic (*mano-mayakosa*) and beautiful (*ananda-mayakosa*) (Tandon, 1995).

Mandukya Upanishad describes four states of consciousness: *Jagrat* (awake state), *Swapna* (dream state), *Sushupti* (deep sleep state) and *Thuriya avastha* (the fourth state – non-dual or thoughtless state). The fourth state is described as a state in which thoughts cease and *mind is still*. There will be a feeling of great happiness or bliss (*Ananda*) and there will be no external or internal impulses reaching from peripheral sensory systems. Patanjali, the father of yoga, in his *yoga sutra* describes fourth state as the one in which "the individual is firmly established in his original state. In all other states, the individual identifies himself with his mind and thoughts". That is, in fourth state, an individual eliminates the limitations by his body, senses, mind, thought and reaches a thoughtless and infinite state of pure awareness (Ramamurhti, 1995).

Though Upanishads and other treatises give different descriptions of stages of consciousness, all of them emphasize on the cosmic consciousness as the source of all consciousness. Also, Indian and other Eastern thoughts emphasize that one can achieve a healthy ASC through practice, different from pathological or physiological ASC (Tandon, 1995). In yoga, the individual is trained to prevent sensory impulses from reaching consciousness – initially both exteroceptive and proprioceptive and later, even the nociceptive. These impulses are not only prevented from reaching consciousness but also from eliciting limbic and autonomic responses. As the yogi practices, fewer and fewer sensory impulses reach consciousness. The Yogi gradually starts controlling his thoughts and consciousness. Two techniques have been described in Hindu-Buddhist traditions for controlling

thoughts. One is to actively suppress the thoughts, which arise and concentrate on only one thought. The second technique is the standing back technique and is the basis of Buddhist Vipasana meditation in which the individual is trained to allow every thought that arises to flow through without in anyway controlling them. In the course of time, the thoughts subside and only one thought will remain. This too disappears after some more time and the yogi reaches the thoughtless fourth state. This is also described in Bhagavadgita as 'realizing that natural elements like mental waves and disturbances reside in the nature of our body, the wise man does not get attached to these and stands back' (Deshmukh, 2009).

Though the possibility of attaining higher states of consciousness was described much earlier, interest resurged in the last few decades with the scientific investigation of Transcendental meditation and Zen traditional meditation (Wallace, 1970). In a landmark study, those who practiced Transcendental meditation (TM) had increased skin resistance, decreased heart rate and oxygen consumption. In addition, there were interesting EEG changes, namely large amplitude alpha waves that were different from sleep (Wallace, 1970). Similar changes were observed in practitioners of Zen meditation (Kasamatsu and Hirai, 1966). The authors interpreted the EEG pattern as special state of consciousness in which the cortical excitatory level becomes lower than the ordinary wakefulness but is not lowered as in sleep and yet outer and inner stimulus is precisely perceived with steady responsiveness. The altered conscious state achieved is also different from those induced by hypnosis or autosuggestion.

A recent study investigated structural changes in brain following meditation. Cortical thickness of a group of mindfulness meditation practitioners and matched controls who were non-practitioners was compared. Interestingly, meditation practitioners had significantly higher cortical thickness in Right anterior insula, right middle and superior frontal sulci than non-practitioners. The thickness was more pronounced in older participants and authors inferred that meditation might offset age-related cortical thinning. In addition, there was a correlation between meditation experience and cortical thickness indicating cortical plasticity is associated with meditation (Lazar et al., 2005).

Conscious–Nonconscious interface

Several physiological and pathological processes affect consciousness. Net effect is a behaviorally observed altered state of consciousness. The external/outward presentation may not necessarily depict the true consciousness. That the consciousness so also its altered state can be experienced is strange but true. The

experience or 'inner' consciousness is poorly examined. Much of this is a teaching from Eastern philosophies.

In Indian Vedanta literature, the self (*purusha*) is described as threefold – the outer *bahyatman*, the inner *antaratman* and the transcendental *paramatman*. The *bahyatman* is the outer body which is made of sense organs. The *antaratman* is the one that perceives the physical and mental states, in other words the consciousness. The transcendental state, *paramatman*, is realized by one who masters the *yoga* (self integration), the *pranayama* (quiet, conscious breathing), *pratyahara* (detached attitude) and *Samadhi* (absolute ultimate reality) (Deshmukh, 2009). This, though an altered state, is a heightened state of consciousness in which the person is beyond senses which is unthinkable and indescribable. The Indian thought emphasizes the continuum along the conscious and non-conscious states and importantly this altered state can be produced willfully.

Attempts have been successfully made to autogenously alter the states of consciousness. While the behavioral manifestations are observable, the experience remains subjective. The efforts have been directed to objectively assess this inner state. Such attempts may be compared to the 'blind men describing the elephant'. Manipulations in the inner consciousness by will have also led to changes in the observed consciousness. The observations included several variables that also measure physiological changes in different organs and brain. The evidence for this autogeneously induced altered states of consciousness has been substantial; neurophysiological changes both at periphery and central are willfully produced in the process of altering the state of consciousness.

Conclusion

Consciousness is an active and dynamic state. The experience or 'inner' consciousness is poorly examined. Much of this is a teaching from Eastern philosophies. Attempts have been successfully made to autogenously alter the states of consciousness. This is a paradigm shift in the past three decades as regards our understanding of mind and matter. If properly harnessed it has the potential for several applications in health sciences.

References

- Brown, R. P. and Gerbarg, P. L. (2005). Sudarshan Kriya Yogic breathing in the treatment of stress, anxiety, and depression. Part II – Clinical applications and guidelines. *Journal of Alternate and Complementary Medicine*, 11, 711–717.
- Cavanna, A. E. and Trimble, M. R. (2006). The precuneus: A review of its functional anatomy and behavioural correlates. *Brain*, 129, 564–583.
- Critchley, H. D., Melmed, R. N., Featherstone, E., Mathias, C. J. and Dolan R. J. (2001). Brain activity during biofeedback relaxation: A functional neuroimaging investigation. *Brain*, 124, 1003–1012.
- Deshmukh, V. D. (2009). *Ipseity, Atman and Consciousness*. New Delhi, D.K. Print world.
- Farthing, G. W. (1992). *The psychology of consciousness*. NJ: Englewood Cliffs.
- Fenwick, P. B., Donaldson, S., Gillis, L., Bushman, J., Fenton, G. W., Perry, I., Tilsley, C. and Serafinowicz, H. (1977). Metabolic and EEG changes during transcendental meditation: An explanation. *Biological Psychology*, 5, 101–118.
- Herzog, H., Lele, V. R., Kuwert, T., Langen, K. J., Rota Kops, E. and Feinendegen, L. E. (1990). Changed pattern of regional glucose metabolism during yoga meditative relaxation. *Neuropsychobiology* 23, 182–187.
- Hofle, N., Paus, T., Reutens, D., Fiset, P., Gotman, J., Evans, A. C. and Jones, B. E. (1997). Regional cerebral blood flow changes as a function of delta and spindle activity during slow wave sleep in humans. *Journal of Neuroscience*, 17, 4800–4808.
- Janakiramaiah, N., Gangadhar, B. N., Naga Venkatesha Murthy, P. J., Harish, M. G., Subbakrishna, D. K. and Vedamurthachar, A. (2000). Antidepressant efficacy of Sudarshan Kriya Yoga (SKY) in melancholia: A randomized comparison with electroconvulsive therapy (ECT) and imipramine. *Journal of Affective Disorders*, 57, 255–259.
- Kalyani, B. G., Venkatasubramanian, G. and Gangadhar, B. N. (2009). Neuro-hemodynamic Correlates of Om Chanting: A fMRI Study. Personal communication.
- Kasamatsu, A. and Hirai, T. (1966). An electroencephalographic study on the zen meditation (Zazen). *Folia Psychiatrica Neurologica Japonica*, 20, 315–336.
- Katayama, H., Gianotti, L. R., Isotani, T., Faber, P. L., Sasada, K., Kinoshita, T. and Lehmann, D. (2007). Classes of multichannel EEG microstates in light and deep hypnotic conditions. *Brain Topography*, 20, 7–14.
- Laureys, S. (2005). The neural correlate of (un)awareness: Lessons from the vegetative state. *Trends in Cognitive Sciences*, 9, 556–559.
- Lazar, S. W., Kerr, C. E., Wasserman, R. H., Gray, J. R., Greve, D. N., Treadway, M. T., McGarvey, M., Quinn, B. T., Dusek, J. A., Benson, H., Rauch, S. L., Moore, C. I. and Fischl, B. (2005). Meditation experience is associated with increased cortical thickness. *Neuroreport*, 16, 1893–1897.
- Loewenstein, R. J. and Putnam, F. R. (2005). Dissociative disorders. *Comprehensive Textbook of Psychiatry*. Editors: B. J. Saddock and V. A. Saddock (pp. 1844–1901). Philadelphia: Lipincott Williams and Wilkins.
- Merskey, H. and Piper, Jr., A. (1998). Treatment of dissociative identity disorder. *American Journal of Psychiatry*, 155, 1462–1473.
- Moruzzi, G. and Magoun, H. W. (1949). Brain stem reticular formation and activation of the EEG. *Electroencephalography and Clinical Neurophysiology*, 1, 455–473.

- Newberg, A. B. and Iversen, J. (2003). The neural basis of the complex mental task of meditation: Neurotransmitter and neurochemical considerations. *Medical Hypotheses*, 61, 282–291.
- Ramamurhti, B. (1995). Fourth state of consciousness Decade of the Brain. Rockville: National Institute of Mental Health.
- Shapiro, S. and Lehrer, P. M. (1980). Psychophysiological effects of autogenic training and progressive relaxation. *Biofeedback and Self Regulation*, 5, 249–255.
- Tandon, P. N. (1995). Concepts of Consciousness: Brain–Mind relationship. Rockville: National Institute of Mental Health.
- Telles, S., Nagarathna, R., Nagendra, H. R. and Desiraju, T. (1994). Alteration in auditory middle latency evoked potentials during meditation on a meaningful symbol ‘OM’. *International Journal of Neuroscience*, 76, 87–93.
- Vaitl, D., Birbaumer, N., Gruzelier, J., Jamieson, G. A., Kotchoubey, B., Kübler, A., Lehmann, D., Miltner, W. H., Ott, U., Pütz, P., Sammer, G., Strauch, I., Strehl, U., Wackermann, J. and Weiss, T. (2005). Psychobiology of altered states of consciousness. *Psychological Bulletin*, 131, 98–127.
- Wallace, R. K. (1970). Physiological effects of transcendental meditation. *Science*, 167, 1751–1754.

Beyond the boundaries of the brain

Dean Radin

Does telepathy – a purported means of communication unmediated by the ordinary senses – exist? Experimental evidence, meta-analyses and debates suggesting an affirmative answer can be found in a growing number of main-stream journals, and a physical basis for telepathy appears to be increasingly plausible based upon considerations of quantum holism and advancements in quantum information processing. Future neuroscience models of conscious and unconscious perception may be obliged to take into account a common human experience once dismissed as mere superstition.

Most people, most of the time, take for granted that *reality* – the objective world “out there,” independent of ourselves – is revealed by common sense. This is as true for the general public as it is for most scientists. Common sense presents a reality where objects are separated from one another in space and time, the world is stable and independent of observers, causality is a mechanistic, deterministic process, time is strictly unidirectional, and so on.

The problem is that we know this picture of reality is false. Not just mildly wrong, but incorrect in fundamental ways. By probing ever deeper into the fabric of reality, physics has revealed that naïve realism is based upon a miniscule slice of a much larger and exceedingly variegated world. Einstein's theories tell us that what were once thought to be basic building blocks of reality – space, time, matter and energy – are not foundational after all, but intertwined relationships. Quantum mechanics shows that objects are not as separate as they appear to be, but entangled beyond the usual boundaries of space and time. Even our most cherished assumptions about local reality – that the world exists independently of observation – has been empirically falsified (Gröblacher et al., 2007). And some physicists now suggest, in the pages of *Nature* no less, that “The only reality is mind and observations, but observations are not of things. To see the Universe as it really is, we must abandon our tendency to conceptualize observations as things” (Henry, 2005).

If exotic physics seems too far removed from human experience, then consider that conscious awareness is the only direct link to reality available to us, and awareness is a mere trickle of the torrents of information unconsciously being processed by the nervous system and brain. In the process of throttling down this information for the sake of cognitive efficiency, the brain tricks us into perceiving not the world as it is, but as we expect it to be, or what might be expedient for the sake of survival. The lesson is that when it comes to pondering the boundaries between objective and subjective realities, it is useful to cultivate tolerance for the unexpected.

And for good reason, for when probing that enigmatic boundary, we can quickly encounter phenomena that aggressively violate common sense. Such experiences tend to be labeled anomalous, paradoxical, paranormal, or if the mismatch is too intense, bizarre or crazy. One such phenomenon, telepathy, involves a class of human experiences documented throughout history, in all cultures, and at all educational levels. These experiences are readily accepted by the majority of the world's population, but the possibility that it is what it appears to be is either studiously ignored in academia, or its mere existence is flatly and occasionally vehemently denied.

A forgotten episode in the origins of the neurosciences

An introspective country youth decided to become an astronomer. While at the university, the pace of big city life disagreed with his simple ways, and so he left school. It being peacetime, he decided to enlist in the cavalry for a year of service, looking forward to a year of riding horses and enjoying the outdoors in relative serenity.

One morning, while on a training exercise, his horse was spooked and suddenly reared. He was tossed into the air and landed hard on the road directly in the path of a fast-approaching, horse-drawn cannon. With sudden horror, he knew that he was about to be crushed. Miraculously, the driver of the artillery battery managed to stop the horses just in time. The accident left him shaken but without serious injury.

At that very moment, a hundred miles away in his family's home, the young man's sister was suddenly overwhelmed with an ominous certainty that something bad had happened to her brother. She anxiously insisted that their father contact him, and he did so using the fastest means possible at that time – a telegram.

That evening, the young man received the telegram and became concerned, because he had never before received a telegram from his father. Then, upon reading his sister's urgent concern about his well-being, he realized that his feelings of

intense fear earlier in the day had somehow reached his sister. Many years later he wrote in his memoirs, “This [was] a case of spontaneous telepathy in which at a time of mortal danger, and as I contemplated certain death, I transmitted my thoughts, while my sister, who was particularly close to me, acted as the receiver” (Berger, 1940).

This experience profoundly transformed the young man’s interests from the mysteries of outer space to the mysteries of inner space. After finishing his military service, he returned to the university and focused on learning medicine, determined to understand how the “psychic energy,” as he called it, could have possibly carried a telepathic message to his sister a hundred miles away.

After many years of concentrated effort, this man finally published his recordings of human brainwaves (in 1929). With this invention, German psychiatrist Hans Berger established his place in the history of the neurosciences, because for the first time he showed that electrical activity of the human brain was correlated with subjective states of mind.

What is less well known is that Berger also pursued an experimental program investigating telepathy in hundreds of people. He was attempting to develop an explanation for his sister’s telepathic episode based upon a “mental radio” analogy. To his regret the analogy proved to be inadequate because it did not explain how his sister could have detected ultra-weak electromagnetic signals over many miles, and even if she could, there were no known means of “decoding” those signals. Berger was stumped by his sister’s apparent “non-local” connection with him, a connection that defied all known physical laws of the day. Not surprisingly, his colleagues did not think highly of his interests, and today hardly anyone remembers the origins of the EEG.

Experimental tests

Many scientists have heard stories from trusted friends that resemble Hans Berger’s experience, or they have personally encountered episodes of their own. Starting in the mid-1800s, controlled experiments were devised to explore whether these effects were best understood in terms of coincidence, misreporting or confabulation, or whether it might be a phenomenon involving genuine distant communication. Over the past half century, one of those methods has been refined and repeatedly tempered into a simple but effective laboratory demonstration.

The method is known as the *ganzfeld* (German for “whole field”) telepathy experiment. A participant in the *ganzfeld* condition reclines in an easy chair in a sound-proof room with translucent hemispheres placed over his or her eyes. The participant’s face is illuminated with a soft red light, the eyes remain open, and

white noise is played over headphones. After a few minutes in this state, most participants accommodate to the uniform stimuli and soon afterwards begin to experience mild visual and auditory hallucinations.

The experiment involves two people, a nominal “sender” and “receiver.” The receiver relaxes in the ganzfeld condition in an isolated, shielded location while a sender, sequestered in a distant room, views a video clip randomly selected out of a large pool of potential target clips. The sender is asked to mentally send the contents of the target clip to the receiver, and the receiver is asked to speak aloud his or her mental impressions into an audio recorder. A typical test session lasts about 30 minutes, during which time the sender might view 15 to 20 repeated presentations of the same video clip.

After the sending period ends, the receiver is taken out of the ganzfeld condition by an experimenter who is also blind to the actual target. The recorded mentation is played back while the participant views four video clips in random order, one the actual target along with three decoys. The clips in these target pools are previously arranged to be as different from one another as possible to help the receiver select just one of the four clips as the best match to his or her subjective impressions.

Under the null hypothesis, the probability that the receiver will correctly guess the actual target is 1 in 4, for a 25% chance hit rate. Over repeated sessions if the average hit rate is sufficiently above 25%, this provides evidence for a communications anomaly that confirms the essence of anecdotal reports of telepathy.

To provide valid interpretation of data collected in these experiments, diligence must be exercised to prevent any conventional means of communication between sender and receiver. Surveillance may be employed to detect subject or experimenter fraud, methods used to ensure adequate randomization of targets, and laboratories obsessively checked to shield against potential cues carried by third party confederates, or by sounds, vibrations, odor, lights, or electromagnetic signals. Most modern versions of these tests use fully automated digital recording and video presentations.

Cumulative results

From 1974 to 2008, 32 principal investigators, representing a dozen laboratories located in Europe, North and South America, published 112 articles describing ganzfeld telepathy experiments. Cumulatively they report 1,498 hits in 4,705 trials, resulting in an overall hit rate of 31.8%, as compared to the chance expected 25%. This hit rate is associated with a p value of 3×10^{-30} .

Starting with Hyman (1985), the accumulated data began to be reviewed using meta-analytic techniques; this history is reviewed in Storm and Ertel (2001) and Radin (2006). Of the half-dozen published reviews, including two by scientists with long histories of skepticism about this phenomenon, all agreed that the effect size was significantly positive. An initially negative conclusion by Milton and Wiseman (1999) was later found to have used inappropriate statistics; when corrected it too revealed a significant outcome. All of the reviewers also agreed that the results could not be plausibly explained by selective reporting practices, and all but one found that effect sizes did not vary with experimental quality. In the one exception, Hyman (1985) claimed that effect sizes declined with improved experimental quality, implying that apparently positive results were due to design or implementation flaws. But ten independent scientists, including two statisticians and two psychologists not previously involved in the debate, reviewed his argument and none agreed with his conclusion. In addition, a team of avowedly skeptical researchers led by Delgado-Romero and Howard (2005) successfully repeated the experiment, obtaining the same 32% hit rate estimated by the meta-analyses.

Meta-analysis of conceptually similar experiments based on conscious awareness of the feeling of being stared at by a distant person, reported by Sheldrake (2003) and others, show similar results (Radin, 2005). In 60 reported experiments comprising a total of 33,357 individual trials, an overall hit rate of 54% was obtained where chance expectation was 50%. The weighted mean effect size for this 4% excess was associated with $p = 5 \times 10^{-17}$.

Neuroscience perspective

If interpretation of these conscious-report experiments is correct, then one might expect to find unconscious physiological correlates between isolated dyads, one of whom is exposed to stimuli at random times while the responses of the other is monitored. Experiments investigating such correlations using EEG measures were first reported in student-teacher dyads by Tart (1963) and in identical twins by Duane and Behrendt (1965), the latter report published in *Science*. Those articles stimulated ten replications by eight groups around the world. Of the ten, nine reported positive results, including reports by Targ and Puthoff (1974) in *Nature*, by Rebert and Turner (1974) in *Behavioral Neuropsychiatry*, and by Orme-Johnson et al. (1982) in the *International Journal of Neuroscience*.

In the 1990s, a new series of EEG correlation studies were reported by Grinberg-Zylberbaum et al. (1994) in *Physics Essays*, again stimulating a flurry of successful replications (including Radin, 2004; Wackermann et al., 2003; and

Standish et al., 2004). To investigate the source of these EEG correlations, experiments conducted by Standish et al (2003) using an fMRI found that the visual cortex of the receiver was significantly activated while the distant sender viewed a flickering checkerboard pattern. This group later successfully replicated their original finding (Richards et al., 2005).

In addition, some three dozen studies conducted by laboratories in the United States, Germany, and the United Kingdom have reported similar effects using autonomic measures, such as skin conductance and heart rate. Meta-analysis of those studies by Schmidt et al. (2004) concluded that they showed significant, repeatable effects, and that alternatives such as selective reporting were insufficient to explain the results.

Discussion

Innumerable anecdotes involving apparent telepathic connections have motivated scientists since the late 1800s to experimentally investigate these experiences. As of 2009, over 200 laboratory experiments studying conscious and unconscious forms of telepathy have been published. Based upon the peer-reviewed empirical evidence, it is clear that distant forms of communication consistent with anecdotal reports of telepathy are repeatedly observed under scientifically controlled conditions. One might expect that this phenomenon would be well accepted by now. But many scientists remain skeptical. Why?

The short answer is that many scientists, perhaps most, assume that what is possible in human experience can only be based on common sense models of reality. From that perspective, telepathy is impossible because it violates the tenets of naïve realism, and thus any positive evidence has to be flawed. There is no alternative, and from this stance, ESP can only mean “error some place” rather than “extra-sensory perception.”

But knowing that naïve realism is an incorrect ontology, we might ask whether a slightly more comprehensive understanding of the fabric of reality would provide more plausible support for telepathy. An affirmative answer seems increasingly likely given that quantum entanglement, or what Einstein called “spooky action at a distance,” has been unequivocally demonstrated as fact in the microscopic world, and a growing number of physicists now believe entanglement also extends into the macroscopic world. As Vedral (2008) writes in *Nature*, “Traditionally, entanglement was considered to be a quirk of microscopic objects that defied a common-sense explanation. Now, however, entanglement is recognized to be ubiquitous and robust.” From this perspective, which reframes the meaning of a physical object, the brain is not completely locked inside the skull. Parts of

it also exist beyond conventional space-time. This suggests that brain processes responsible for ordinary perception may be influenced not only by local sensory information, but by impressions received from what common sense would label “at a distance.” Evidence supporting the existence of such effects in living systems can now be found in studies involving photosynthesis at room temperature, which “contrasts with the long-held view that long-range quantum coherence between molecules cannot be sustained in complex biological systems, even at low temperatures” (Collini et al., 2010).

Could the brains of Hans Berger and his sister have been entangled not just emotionally, but physically? Could those brains have “co-processed” information regarded as important to both individuals? Does this “psychophysical entanglement” lead to remarkable coincidences between separated identical twins, such as a case reviewed by Playfair (2003) in which a healthy man suddenly experienced a suffocating pain in his chest while hundreds of miles away, at the same moment, his identical twin suffered a heart attack? At this point such ideas are purely speculative, but such possibilities are no longer unthinkable.

In sum, telepathy is a common human experience once dismissed as mere superstition. But faced with countless anecdotes, accumulating empirical confirmation, and rising theoretical support, it appears to merit a second look. The implications of genuine telepathy for our understanding of perception, conscious and unconscious processes, the sense of personal self, and many other aspects of subjective experience, are intriguing.

References

- Berger, H. (1940). *Psyche*, p. 6. Jena: Gustav Fischer.
- Collini, E., Wong, C. Y., Wilk, K. E., Curmi, P. M. G., Brumer, P. and Scholes, G. D. (2010). Coherently wired light-harvesting in photosynthetic marine algae at ambient temperature. *Nature*, 463, 664–648.
- Delgado-Romero, E. and Howard, G. (2005). Finding and correcting flawed research literatures. *The Humanistic Psychologist*, 33, 293–303.
- Duane, T. D. and Behrendt, T. (1965). Extrasensory electroencephalographic induction between identical twins. *Science*, 150, 367.
- Grinberg-Zylberbaum, J., Delaflor, M., Attie, L. and Goswami, L. (1994). The Einstein-Podolsky-Rosen paradox in the brain: The transferred potential. *Physics Essays*, 7, 422–428.
- Gröblacher, S., Paterek, T., Kaltenbaek, R., Brukner, Č., Żukowski, M., Aspelmeyer, M. and Zeilinger, A. (2007). An experimental test of non-local realism. *Nature*, 446, 871–875.
- Henry, R. C. (2005). The mental universe. *Nature*, 436, 29.
- Hyman, R. (1985). The ganzfeld psi experiment: A critical appraisal. *Journal of Parapsychology*, 49, 3–49.

- Milton, J. and Wiseman, R. (1999). Does Psi exist? Lack of replication of an anomalous Process of information transfer. *Psychological Bulletin*, 125, 387–391.
- Orme-Johnson, D. W., Dillbeck, M. C., Wallace, R. K. and Landrith, G. S. (1982). Intersubject EEG coherence: Is consciousness a field? *International Journal of Neuroscience*, 16, 203–209.
- Playfair, G. L. (2003). *Twin Telepathy: The Psychic Connection*. London: Vega.
- Radin, D. I. (2004). Event-related EEG correlations between isolated human subjects. *Journal of Alternative and Complementary Medicine*, 10, 315–324.
- Radin, D. I. (2005). The sense of being stared at: A preliminary meta-analysis. *Journal of Consciousness Studies*, 12 (6), 95–100.
- Radin, D. I. (2006). *Entangled Minds*. New York: Simon & Schuster.
- Rebert, C. S. and Turner, A. (1974). EEG spectrum analysis techniques applied to the problem of psi phenomena. *Behavioral Neuropsychiatry*, 6, 18–24.
- Richards, T. L., Kozak, L., Johnson, L. C. and Standish, L. J. (2005). Replicable functional magnetic resonance imaging evidence of correlated brain signals between physically and sensory isolated subjects. *Journal of Alternative and Complementary Medicine*, 11, 955–963.
- Schmidt, S., Schneider, R., Utts, J. and Walach, H. (2004). Distant intentionality and the feeling of being stared at: Two meta-analyses. *British Journal of Psychology*, 95, 235–247.
- Sheldrake, R. (2003). *The sense of being stared at*. New York: Crown Publishers.
- Standish, L. J., Johnson, L. C., Richards, T. and Kozak, L. (2003). Evidence of Correlated functional MRI signals between distant human brains. *Alternative Therapies in Health and Medicine*, 9, 122–128.
- Standish, L. J., Kozak, L., Johnson, L. C. and Richards, T. (2004). Electroencephalographic evidence of correlated event-related signals between the brains of spatially and sensory isolated human subjects. *Journal of Alternative and Complementary Medicine*, 10, 307–314.
- Storm, L. and Ertel, S. (2001). Does psi exist? Comments on Milton and Wiseman's (1999) meta-analysis of ganzfeld research. *Psychological Bulletin*, 127, 424–433.
- Targ, R. and Puthoff, H. (1974). Information transmission under conditions of sensory shielding. *Nature*, 252, 602–607.
- Tart, C. T. (1963). Possible physiological correlates of psi cognition. *International Journal of Parapsychology*, 5, 375–386.
- Vedral, V. (2008). Quantifying entanglement in macroscopic systems. *Nature*, 453, 1004–1007.
- Wackermann, J., Seiter, C., Keibel, H. and Walach, H. (2003). Correlations between spatially separated human subjects. *Neuroscience Letters*, 336, 60–64.

Plants of the gods and shamanic journeys

Elaine K. Perry and Valerie Laws

Among agents which alter the boundary between conscious and non conscious cognition, the ritualistic use of plant species (often in a spiritual context, hence 'plants of the gods') provides an example of long-standing empirical knowledge subsequently verified by scientific (chemical, pharmacological and psychological) evidence. Based on such an impressive record of acquired knowledge, exploration of experiences of the shaman, who deliberately enters an altered state of consciousness to obtain otherwise inaccessible information allegedly from other 'dimensions' of consciousness, may contribute new insights in the neuroscience of consciousness.

Inducing altered conscious awareness

'Plants of the gods' provide a fascinating example of an interface between empirical and scientific knowledge providing insights into novel aspects of conscious non conscious interactions. Hundreds of ethnobotanicals (traditionally used plants) with potent neuroactivities have been used ritualistically for thousands of years to explore conscious experience, inducing altered states interpreted in a spiritual or religious context (Hofmann et al., 2001; Perry, 2002). The active plant chemicals are mainly alkaloids interacting with specific transmitter systems in the brain. Alterations in consciousness include hallucinations or visions, experienced in deliberately induced, euphoric trance-like states, described as mystical, inspiring, enlightening, with self dissolution or merging with a 'global' consciousness.

The German psychiatrist Emil Kraepelin considered that the use of pharmacological agents to explore consciousness was an important scientific investigation although primarily as a way of understanding mental disease – reminiscent of divine madness or dionysiac frenzy. A century later the use of chemicals to explore and expand conscious experience continues to intrigue. According to Paul Devereux our modern Western culture is distinct in human history because of the difficulty in accepting the role of natural, mainly plant, chemicals in 'aiding

mind expansion'. While past research focused on psychedelic effects as models of psychiatric disorders, in the last decade neuroscientific studies (brain imaging or cognitive assessment for example) have focused more on how these agents inform brain/ mind mechanisms.

This chapter explores divergent hypotheses that such changes in conscious awareness reflect chemically induced system changes in the brain which alter the selection/emergence into subjective conscious awareness of information. This can be from either: non conscious processing of perceptions, memories and related cognitive functions dependent on neural circuitry; or additional perceptions derived non locally from 'beyond the brain', a concept discussed below. The first is consistent with current neuroscientific experimental knowledge and the second with mystical/shamanic experience and understanding, each with dramatically distinct implications for theories of consciousness based either solely within, or additionally beyond, the brain.

Species, states and neuropharmacology

The following account of ritualistically used plant species, active chemicals, transmitter interactions and subjective experiences is derived from two sources: culturally, or ethnographically based accounts of an individual's subjective experience of altered consciousness; and objective chemical and pharmacological information on the active plant constituents, how these mimic or block the action of transmitters and affect network activities (original references not listed in this chapter are to be found in a previous chapter on plants of the gods, Perry, 2002). There is a dichotomy in values between these sources of information, with cultural accounts focussing on perceived enhancement of mental experience, and much of the pharmacology on negative aspects such as induction of psychosis, toxicity or addiction. Given the diversity of plant species, complexity of chemical constituents and polypharmacology, together with the individuality of subjective experiences and responses to exogenous chemicals, Table 1 is necessarily a highly simplified summary of some major features.

Table 1. Psychedelic plants commonly used in shamanism:
Botanical, chemical, pharmacological and psychological characteristics

Plant species	Active chemicals identified	Receptor action (if known)	Target brain areas/effects (if known)	Subjective experiences
Psilocybe fungi / magic mushrooms e.g. yellow caps <i>Psilocybe cubensis</i>	Psilocybine or psilocyne	Preferential 5HT _{2A} receptor agonist	5HT ₂ receptors concentrated in cerebral cortex, especially visual	Dissolution boundary between concepts of self and non-self and significant/meaningful 'mystical' experiences
Morning glory <i>Ipomoea</i> species	Ergoline derivatives which are lysergic acid like alkaloids; precise active chemical not identified	5HT _{2A} receptor agonists, similar to LSD	Presumed as for above	Visual perceptual changes including hallucinations
<i>Virola</i>	Indoleamines based on 5HT eg 5-MeO-DMT (N, N-dimethyl-tryptamine), 5-OH-DMT (Bufotenine), and also N,N-DMT (most powerful DMT derivative)	Broad spectrum 5HT receptor agonist	Widespread	Hallucinations and also nightmares; some similarities to effects of ayahuasca
Ayahuasca <i>Banisteriopsis</i> spp. vine, with added leaves of shrubs from the <i>Psychotria</i> genus	Complex mixture beta carboline alkaloids eg harmine and tetrahydroharmine, with additives containing DMT	DMT is both a 5HT ₂ and sigma ₁ (voltage sensitive ion channel) receptor agonist; monoamine oxidase inhibitor activities prevent its breakdown	Increases blood flow frontal cortex and paralimbic regions; increases activity (functional imaging) heteromodal association cortex and limbic structures; increases global gamma coherence; sigma ₁ receptors concentrated in striatum, thalamus and frontal cortex	Enhanced introspective attention, altered somatic perceptions and intense emotional modifications, frequently accompanied by visual imagery; transcendental descriptions e.g., 'soul' leaving body and 'encounters with aliens'

Table 1. (continued)

Plant species	Active chemicals identified	Receptor action (if known)	Target brain areas/effects (if known)	Subjective experiences
Peyote cactus <i>Lophophora Williamsii</i>	Mescaline and also low levels of ecstasy-like phenethylamines	5HT receptor (multiple subtypes) agonists or partial agonists	5HT1A concentrated in limbic areas; 5HT1B in visual cortex	Novel, highly pleasurable, mainly visual and often intensely coloured images; conviction of reality
Solanaceae species such as mandrake (<i>mandragora officinalis</i>), <i>atropa belladonna</i> , henbane (<i>hyoscamus niger</i>), <i>datura</i> (e.g. <i>Datura stramonium</i>)	Tropane alkaloids including atropine, scopolamine, hyoscamine amongst many others	Muscarinic cholinergic receptor (broad spectrum subtypes) antagonists	Cerebral cortex (particularly visual) and striatum.	Visual imagery, particularly of familiar people and animals, often nightmare like, and with sensation of flying
Fly agaric <i>Amanita muscaria</i>	Muscimol, ibotenic acid and (less involved in hallucinogenic effect) muscarine	GABAA and glutamate NMDA receptor agonists	Widespread	Hallucinations ranging from ecstatic to demonic
Cannabis <i>Cannabis sativa</i>	Cannabinoid terpenes, Tetrahydrocannabinol (THC) being the most psychoactive	Cannabinoid CB1 receptor agonist	CB1 concentrated in cerebral cortex, hippocampus and amygdala	Euphoria and enhanced perception
Opium poppy <i>Papaver somniferum</i>	Alkaloids such as morphine, codeine, and papaverine	Opioid receptor agonists; μ subtype related to euphoria	Concentrated in cerebral cortex	Waking dream like state with euphoria
Divine sage <i>Salvia divinorum</i>	Diterpene, salvinorum A	Opiate receptor κ subtype	Associated with reward pathways.	Altered perception including hallucinations; said to be safest and mildest of shamanic plants

Neurotransmitter mechanisms

A broad spectrum of transmitter systems are targeted by the chemicals in these plant species, although serotonergic and cholinergic activities are the most prevalent. Amongst species described by Hoffman et al. (2001) 90% of those with known active constituents contain chemicals interacting with these two neurotransmitter systems. Our understanding of the pharmacological mechanisms is still incomplete. Potent interactions of DMT (see Table 1) with the sigma 1 (non opioid) receptor have recently been identified (Su et al., 2009). Sigma receptors (known as orphan receptors due to unknown brain functions) regulating a variety of signal transductions, are highly expressed in deeper laminae of the cortex, and have been implicated in memory and drug dependence (Hayashi and Su, 2005).

A broad classification of 'plants of the gods', based on correlated subjective experiences and neuropharmacological mechanisms (receptor targets and brain areas affected), has been proposed as follows (Rogers and Perry, in preparation): *Mild* – euphoric / dreamy / enhanced visual perception associated with activation of opiate receptors (μ subtype – cerebral cortex, thalamus and ventral tegmental area, or κ subtype – parietal and temporal cortex, nucleus accumbens, amygdala and thalamus) or the cannabinoid receptor CB1 (cerebral cortex and hippocampus); *Marvellous* – sensory, especially visual, intensification both pleasurable and meaningful, associated with activation of broad spectrum serotonin receptor subtypes (5HT1B concentrated visual cortex and 5HT1A in limbic areas); *Mystical* – dissociation / depersonalisation / 'transcendence' associated with activation of the serotonin receptor 5HT2A activation (highest in frontal, parietal and somatosensory cortex) and sigma 1 receptor (highest striatum and thalamus, intermediate frontal cortex); and *Monstrous* – nightmarish visions eg animals/people with or without flying sensation associated with broad spectrum muscarinic receptor blockade (M1 being highest in hippocampus, cerebral cortex and striatum) or combined GABA and glutamate receptor activation. This classification is but one of many, and Fischer (1971) proposed a division between ecstatic and meditative altered states: the first along the perception hallucination continuum (creative, psychotic, ecstatic) and the second along the perception meditation continuum (hyperarousal states as in some Buddhist and yogic practices).

Neurochemically, experiences with anti-cholinergic agents, which are potentially highly toxic, include hallucinations that are often negatively experienced. The German toxicologist Gustav Schenk writing in 1955 after experimenting with the fumes from henbane seeds, reported: "There were animals which looked at me keenly with contorted grimaces and staring terrified eyes ... at the same time I experienced an intoxicating sensation of flying ... I soared where my hallucinations – the clouds, herds of beasts were swirling along".

In contrast, agents with serotonin receptor activity, particularly 5HT_{2A}, are associated with fantastic, euphoric, transcendental or mystical experiences. The psychologist, Havelock Ellis's described effects of peyote: "Visions never resemble familiar objects; they were extremely definite, but yet always novel; they were constantly approaching, and yet constantly eluding, the semblance of known things ... glorious fields of jewels, solitary or clustered, sometimes brilliant and sparkling". Aldous Huxley (1954) reported: "This is how one ought to see, how things really are. And yet there were reservations. For if one always saw like that, one would never want to do anything else. Just looking, just being divine." The Mazatec shaman, Maria Sabina described psilocybine effects as: "A world beyond ours, a world that is far away, nearby and invisible. And this is where God lives, where the dead live, the spirits and the saints, a world where everything has already happened and everything is known. That world talks. It has a language of its own. I report what it says."

With respect to 'transcendental', mystical, spiritual, experiences, sometimes referred to as 'psi' effects, plant species targeting serotonin receptors are probably the most relevant. Serotonin receptors play an important role in perception, affecting regulation of attention. Pharmacological challenge with the 5-HT_{2A} agonist psilocybine has been used to study the neurobiological basis of cognition including attention (Vollenweider, 2006). In healthy volunteers a double-blind, placebo-controlled dose-effect study of psilocybine identified increased "general inactivation", "emotional excitability", and "dreaminess" (Hasler et al., 2004). Ayahuasca increases frontal lobe activity (Riba et al., 2006). Serotonin has also been implicated in spiritual/mystical experiences on the basis of similarities between LSD and non drug mystical experiences (Goodman, 2002; Borg et al., 2003). That the brain normally produces low levels of DMT (Su et al., 2009) raises the intriguing question of how individual variations in brain DMT levels, normal or drug induced, alter conscious states.

Models of altered states

One of the key functions of the normal brain is to filter out irrelevant or difficult to process, experiences, so focusing only on useful information. Plant based or synthetic psychedelics generically reduce this censorship, allowing uncensored conscious access to non conscious processing. Psilocybine, for example, reduces the ability to suppress or ignore distracting stimuli but does not reduce attentional capacity (Carter et al., 2005). Ahahuasca promotes EEG gamma synchrony (Stuckey et al., 2005), a physiological state associated with conscious awareness (see LeBeau chapter) which is enhanced in lucid dreaming (Voss et al., 2009; see also Voss chapter).

Mechanisms implicated in hallucinatory experiences include altered cortical inhibition and/or dream intrusion. Specific abnormalities in perception and attention with cholinergic hypoactivity in frontal and temporal cortical areas have been implicated in hallucinations occurring in degenerative brain diseases, similar to those induced by anti-muscarinic agents in normal individuals (Collerton et al., 2005). Cholinergic neurons in the basal forebrain are central to this, and also to dream intrusion models (the idea that psychotic symptoms reflect a dream like state during waking), since these (together with brainstem cholinergic neurons) are more active during waking and REM (rapid-eye movement) sleep than during non-REM sleep (dreaming being more intense in the latter). Effects of antimuscarinic agents to induce hallucinations may resemble hypnagogic or hypnopompic hallucinations which occur at the onset and offset of sleep respectively, and often include images of people or animals. At these points in the sleep cycle cholinergic activity may be transiently more reduced than monoaminergic, including serotonin, activity.

Neurotransmitter correlates of conscious awareness have focused on cholinergic projections to cortex and thalamus “confining the conscious stream” (Perry et al., 1999). Evidence on neuronal correlates of natural variations, drug induced and disease associated changes in consciousness are all consistent with a key role for acetylcholine. Since the serotonergic system is generally inactive during the sleep cycle, and since dreaming involves varying degrees of conscious awareness, a similarly global involvement of 5HT in conscious awareness is not so readily supported. However, hallucinations induced in the awake state by serotonergic agonists (Table 1), and the prevalence of 5HT₂ receptors in the cortex indicate that in the awake state this system is important in controlling conscious non-conscious trafficking.

Based on neuronal and psychological models of altered states of consciousness, most neuroscientists assume that visions induced by chemical hallucinogens reflect access to normally suppressed memories and associations. The shamanic practitioner believes otherwise – that the visions are based on reality not normally accessible to everyday consciousness. In shamanism, the practitioner achieves states of altered consciousness deliberately induced using plant extracts or via other strategies (drumming or dancing for example) to experience ‘soul journeys’ and contact with a ‘spirit’ realm. Aldous Huxley posited that such states, in some instances at least, are actual revelations of realms usually hidden from human conscious awareness.

It might be worth considering whether subjects on shamanic journeys or ‘astral projection’, could be somehow freed from perceiving merely three dimensionally. After all, string theory has postulated there are ten dimensions. We cannot step outside the brain to examine such a possibility. We are ‘of’ three dimensions; mathematical points and lines, supposedly in one and two dimensions, are

actually three dimensional models of notional one and two dimensional concepts. Recounting after shamanic journeying what was 'seen' and experienced, what was perceived 'outside' three dimensional perception, is possibly translated into 'normal' three dimensional terms as metaphorical representations of what was experienced in some multidimensional field, recounted in the only way the normally conscious brain can both hold and express them. We naturally think metaphorically. Perhaps William Blake's challenging and prophetic, 'How do we know, but every bird that cuts the airy way, is an immense world of delight, closed by our senses five?' applies here.

Beyond the brain?

Science relies on so-called objective evidence, obtained in the absence of any preconceived bias or belief by independent experimentation and observation, available to any individual trained in the relevant methodology. The following examples (cited Perry, 2002) provide possible evidence for conscious awareness of information beyond brain activities and cognition dependent on normal sensory experiences: *Michael Harner*, an anthropologist, reported the 'reality' of ayahuasa induced visions of "giant reptilian dragon creatures with pterodactyl-like wings, from whose carnivorous jaws gushed a torrential flood of water, claiming to be masters of the universe that in escaping an enemy created terrestrial life forms"; Harner, troubled by these visions, sought advice from a Conibo Indian shaman who assured him 'they' are always saying this but are only masters of outer darkness; Harner was stunned – he had apparently visited a world well known to the shaman from personal experience; and *Kenneth Kensing* (another anthropologist) was informed at an ayahuasa party, by six out of nine Peruvian Cashinahua present, of the death of his father-in-law which he was later informed by radio had occurred two days before. The first example is surprisingly consistent with an extraordinary hypothesis proposed by Francis Crick: that hallucinogenic drugs like LSD activate portions of 'silent' DNA which contain information about alien beings who 'planted' life on earth in the form of primordial DNA (Crick, 1981).

While many anecdotal reports of paranormal experiences (e.g. telepathy or clairvoyance) induced by psychedelic drugs exist, little or no scientific evidence for such phenomena is available. Two particular 'psi' phenomena are however associated with significant scientific data published in peer reviewed journals: telepathy and psychokinesis (Radin, 1997 and 2006, see also Dean Radin's chapter in this book including reference to over 100 such articles on telepathy published over the last 30 years). Recent studies published in the last few years on telepathy or distant intentionality are summarised in Table 2. Individuals are reported to

Table 2. Scientific evidence supporting the possible existence of non-local or 'psi' phenomena

Phenomenon	Data	Reference
Extra sensory perception: non local observation as a source of knowledge	Significant effects of visualization on a quantum system consisting of a pattern of light produced by interferometer; effects primarily due to a subgroup of experienced meditators	Radin (2008)
Telepathy in connection with e-mail communications	Of 552 trials, 43% guesses were hits, significantly above the chance expectation of 25%	Sheldrake and Smart (2005)
Correlations between distant intentionality and brain function in recipients	Functional magnetic resonance imaging analysis with significant differences between experimental (send) and control (no send) procedures; areas activated – anterior and middle cingulate, precuneus and frontal cortex	Achterberg et al. (2005)
Event-related EEG correlations between isolated human subjects	Experimental test resulted in significant correlation compared to control, with stronger responses associated with larger corresponding responses in non stimulated subjects	Radin (2004)
Distant intentionality and the feeling of being stared at	Meta analysis of 36 studies demonstrating small significant effect of intentions on electrodermal activity of remote subjects	Schmidt et al. (2004)
Correlations between brain electrical activities of two spatially separated human subjects	Simultaneous EEG responses following presentation image to one of separated people with significant differences compared non-stimulation periods	Wackermann et al. (2003)
Qi-gong (a traditional Chinese medicine based on energy fields) healing effects on cell biology	Significantly enhanced phagocytosis, immune function, tumor cell death and reductions in lymphoma growth, in response to the intention of trained practitioners	Six reports published 2001–2003 in American Journal of Chinese Medicine or Journal of Alternative and Complementary Medicine

Based on Medline, 2000–2009 (search terms such as telepathy, clairvoyance, psychokinesis).

affect others or physical objects without any explicable mode of interaction. Such interactions are rare but highly divergent from chance with extremely high levels of statistical significance. If valid (and some of the reports itemised in Table 2 have been followed by correspondence in the same journal, arguing spuriousness due to sample size, effects size and/or publication bias), rare effects are not perhaps surprising since such interactions would potentially involve negative/hazardous interference between individuals and objects. Filters to exclude any such information may have evolved to maintain this at the level of non conscious processing, only rarely accessible via psychedelic drugs or other stimuli. If so, conscious awareness would need to be considered as emerging from non conscious processes of, not only the individual, but also some kind of globally connected state.

Most neuroscientists reject such evidence as spurious, excluding any consideration of implications for understanding or experiencing consciousness. But, by rejecting anomalous data, new insights could be missed. In quantum physics 'reality' was radically revised to include unaccountable but measurable non local connections between apparently isolated objects. Views such as an observer creating reality and consciousness as the most fundamental state/principle, or the existence of parallel universes are commonplace. Such radical concepts stimulating the minds of physicists and non experts alike have yet to be widely considered in the field of neuroscience. Yet many grappling with potential mechanisms of consciousness have reached an impasse on the 'hard' problem, or 'qualia' issue of how subjective experience can be represented or explained in terms of neurons/transmitters/synapses/networks/oscillations. This vacuum will, it is assumed by most, vanish with future advances in neuroscience and, in the absence of consensual information on non local consciousness connections to support any paradigm shift, there are no inspired 'thought experiments' as in quantum physics because of the perceived lack of hard evidence to support any paradigm shift.

In contrast to neuroscientific theories, shamanic practitioners claim there are realms accessible to the trained/tuned mind and that consciousness is a field phenomenon (existing as a global / all encompassing phenomenon independent of the individual). Their belief/perception is of multiple realities/'worlds', similar to physicists' contemporary ideas of multiverses or ten dimensional string theory, that can be explored in expanded states of consciousness and that entities/'spirits' in visions are as real as the physical. If this knowledge is as securely based as scientifically verified shamanic knowledge of the use of plants to alter consciousness (exact species, preparation, dose, effects), it should be worth investigating. If verified, paradigm shifts in consciousness research would occur, with major implications for neuroscientific theories and medical practice.

Those few exploring implications and impacts of 'psi' for theories of consciousness frequently invoke aspects of quantum physics (see also Radin's chapter

in this book on quantum entanglement). Quantum mechanics is notoriously counterintuitive to the extent that it has been claimed few if any physicists fully understand it. The theory is even harder for non-physicists to understand and apply to new fields such as consciousness. Quantum theory, originally developed to explain 'gaps' left by Newtonian physics, has gained respect as explaining the functioning of the universe at subatomic level, with implications for our everyday experience – sunlight leaves the sun through quantum tunnelling for example. Biomedical science has yet to incorporate this kind of new thinking. The principle that the observer inevitably changes the observed, for example, means that the notion of the truly detached independent scientific observer, still required for scientific proof, is challenged.

It has been suggested that the brain uses quantum processes, possibly measured in femtoseconds, as well as currently detectable functioning on a 'slower' scale. Neurones communicate by exchanging packets of chemicals across the synapses, but also by direct electrical spread, with possible parallels with the wave/particle duality of light. Schumann (2007) has suggested a 'metabrain', an overall universal consciousness, with each of our individual 'sub brains' perceiving its own version of reality. He claims that the physical acts on the mental, but also the mental on the physical, and uses quantum models to pursue this thesis. That principles of quantum physics might extend to brain function receives potential support from increasing evidence that quantum phenomena are not just confined to subatomic particles, but extend to the macroscopic including living organisms. Other ideas on consciousness incorporating quantum physical effects include consciousness being a collapse of quantum wave function (Penrose, 1989), particularly applicable to microtubules (Hameroff, 2001), and non local brain interactions involving quantum entanglement (Radin chapter).

Those convinced of the limited, but significant, evidence for psi invoke a range of quantum and other possible mechanisms which however are not readily comprehensible or open to investigation. In one particular traditional Chinese medical practice, the concept of Qi (as an all pervasive energy that can be channelled or re distributed to enhance health) is based on the proposed existence of energy fields within (meridia relevant to acupuncture for example) and around individuals (field effect), that relate to other concepts of 'beyond the brain' activities. A series of reports on significant Qi effects on biological processes, such as in lymphocyte production apparent in strictly controlled experiments, provide limited evidence of non conventional effects of conscious activity. In one study (Fukushima et al., 2001) effects were in part mimicked by microwave or infrared laser pulse treatment. Such heat and light are at least readily comprehensible phenomena from which potential new ideas on how mind extending beyond matter might emerge although just simulating a phenomenon does not necessarily

imply common mechanisms. In addition, not all such studies have been replicated (Yount et al., 2004) indicating the need in this, as in the whole area of non local connectivity, for much more controlled investigation.

Shamanic 'journeying', including experiencing traveling along a tunnel ending in a light space populated by people, animals or other entities communicating meaningful information is reminiscent of reports of near death experiences (NDEs). NDEs and out of body experiences include remote viewing (perceptions not accountable by normal sensory experience), undergone while the subject is deeply unconscious or clinically dead as currently defined. NDEs, particularly remote viewing in an individual with no measurable brain activity, have not yet been verified scientifically as having validity outside the individual's own brain, eg by being perceived by other, conscious, observers such as medical staff. Laws and Perry (2010) suggest that NDEs are potentially easier to attempt to scientifically verify than shamanic journeys, though such verification would open up the possibility of these wider zones of consciousness accessed by shamans. They suggest a first step would be to focus on collecting cases of NDEs, sifting them for accurately reported, externally verifiable, sensory impressions such as sights or sounds, rather than transcendental visions, recounted by 'flatlining' patients (in intensive care) on recovery. If proven, the existence of individual consciousness outside a functioning brain: the possibility of 'seeing' without functioning eyes: would provide the greatest challenge for neuroscience. Ideas of mind-mind or mind-matter non local effects of consciousness may be far reaching, but the idea of an individual being conscious independent of the brain activity is orders of magnitude beyond. An alternative explanation for NDEs, that apparent remote viewing might originate from telepathic/clairvoyant communication with another observer, would need to be excluded for NDEs to be uniquely important amongst 'beyond the brain' phenomena, for theories of consciousness. If brain activity, as currently understood and detected, is not a critical requirement, then investigation of 'psi' phenomena that include measurements of brain activity would not necessarily provide a definitive answer. In a major recent study no significant fMRI imaging effects were measurable in conjunction with reported 'psi' activity (Moulton and Kosslyn, 2008), a finding that was interpreted by the authors as "the strongest evidence yet obtained against the existence of paranormal mental phenomena", rather than raising the possibility that such imaging technology failed to detect correlates of 'psi'. Forms of consciousness may exist, at present undetectable (perhaps relating to quantum processes at subatomic levels or superfast speeds) and with respect to NDEs for example a brain considered 'dead' or not showing any currently detectable function, may function at some as yet unknown level.

Scepticism is as susceptible to wishful and delusional thinking as belief in paranormal phenomena, since both satisfy emotional needs involving our fears and hopes about life and death. Adherents to the 'detached scientific observer' model would like to believe these emotions can be somehow excised, but quantum theory and indeed common sense would tend to refute this. We may not be able to create a phenomenon in a lab, but that does not necessarily mean it is false.

According to the quantum physicist Bernard Haisch (2006) "it is not matter that creates the illusion of consciousness, but consciousness that creates the illusion of matter". Such concepts, however far reaching or incredible to neuroscientists in the face of current evidence, have much in common with the shamanic view of reality that the world is as one dreams: "There is nothing but the spirits of all and everything we see here is something like a shadow from that world" (Black Elk Medicine Man). According to the American psychiatrist Richard Chessick (2009): "We simply do not know enough about the physical world of matter at this time but eventually perhaps discoveries about the nature of matter may solve the problem. But as of now, the implication of this impasse is that the introspective data of consciousness are ontologically subjective, pointing to the absolute necessity for our studying this data in its own domain through introspection". Shamanic exploration of consciousness is one such domain.

While understanding remains evasive, two aspects relating to experiencing consciousness arising from shamanic practises are finally worth mentioning. Several plant agents such as in ayahuasa or peyote, used in a supervised context, are widely reported to have antidepressant effects, reduce panic attacks or addictions. In a follow up study after the use of psilocybine in a controlled trial, the experience was rated among the most personally meaningful and spiritually significant of the participants' lives (Griffiths et al., 2008). More speculatively, but also potentially clinically important, shamanic claims for 'ectopic' influences on mental health ('spirit invasion' or 'soul loss') are so far supported by only occasional scientific reports. Brain pathologies in disorders such as dementia, including cholinergic deficits (Perry and Perry, 1995) with relative serotonergic hyperactivity, are similar to pharmacological effects of many 'plants of the gods' and could, by reducing normal filter mechanisms, expose a minority of individuals to confusing or disturbing experiences of consciousness beyond the brain (Perry, 1995).

In conclusion, the impressive achievements of shamanism in identifying plants that alter the conscious non conscious interface, leading to scientific discoveries of neuroactive plant chemicals and brain mechanisms, suggest that scientific investigation of experience of shamanic consciousness itself could contribute to paradigm shifts in understanding the nature and scope of consciousness.

References

- Achterberg, J., Cooke, K., Richards, T., Standish, L. J., Kozak, L. and Lake, J. (2005). Evidence for correlations between distant intentionality and brain function in recipients: A functional magnetic resonance imaging analysis. *Journal of Alternative and Complementary Medicine*, 11, 965–971.
- Borg, J., Andrée, B., Soderstrom, H. and Farde, L. (2003). The serotonin system and spiritual experiences. *American Journal of Psychiatry*, 160, 1965–1969.
- Carter, O. L., Burr, D. C., Pettigrew, J. D., Wallis, G. M., Hasler, F. and Vollenweider, F. X. (2005). Using psilocybin to investigate the relationship between attention, working memory, and the serotonin 1A and 2A receptors. *Journal of Cognitive Neuroscience*, 17, 1497–1508.
- Chessick, R. D. (2009). Implications of the current insolubility of the mind-brain problem for the contemporary practice of psychodynamic psychiatry. *The Journal of the American Academy of Psychoanalysis and Dynamic Psychiatry*, 37, 315–351.
- Collerton, D., Perry, E. and McKeith, I. (2005). Why people see things that are not there: A novel Perception and Attention Deficit model for recurrent complex visual hallucinations. *Behavioral and Brain Sciences*, 28, 737–757.
- Crick, F. (1981). *Life Itself: Its origin and nature*. New York: Simon & Schuster.
- Fischer, R. (1971). A cartography of the ecstatic and meditative states. *Science*, 174, 4012–4026.
- Fukushima, M., Kataoka, T., Hamada, C. and Matsumoto, M. (2001). Evidence of Qi-gong energy and its biological effect on the enhancement of the phagocytic activity of human polymorphonuclear leukocytes. *American Journal of Clinical Medicine*, 29, 1–16.
- Goodman, N. J. (2002). The serotonergic system and mysticism: could LSD and the nondrug-induced mystical experience share common neural mechanisms? *Psychoactive Drugs*, 34, 263–272.
- Griffiths, R., Richards, W., Johnson, M., McCann, U. and Jesse, R. (2008). Mystical-type experiences occasioned by psilocybin mediate the attribution of personal meaning and spiritual significance 14 months later. *Journal of Psychopharmacology*, 22, 621–632.
- Haisch, B. (2006). *The god theory, universes, zero point fields and what's behind it all*. Waiser, Red Wheel.
- Hameroff, S. (2001). Consciousness, the brain, and spacetime geometry. *Annals of the New York Academy Sciences*, 929, 74–104.
- Hasler, F., Grimberg, U., Benz, M. A., Huber, T. and Vollenweider, F. X. (2004). Acute psychological and physiological effects of psilocybin in healthy humans: A double-blind, placebo-controlled dose-effect study. *Psychopharmacology*, 172, 145–156.
- Hayashi, T. and Su, T. (2005). The sigma receptor: Evolution of the concept in neuropsychopharmacology. *Current Neuropharmacology*, 3, 267–280.
- Huxley, A. (1954). *The Doors of Perception*. London: Chatto and Windus.
- Laws, V. and Rerr, E. (2010). Near death experiences: A new algorithmic approach to verifying consciousness outside the brain. *NeuroQuantology*, 8, 142–154.
- Moulton, S. T. and Kosslyn, S. M. (2008). Using neuroimaging to resolve the psi debate. *Journal of Cognitive Neuroscience*, 20, 182–192.
- Penrose, R. (1989). *The Emperor's New Mind: Concerning computers, minds, and the laws of physics*. Oxford: Oxford University Press.

- Perry, E. (1995). Conscious awareness and neurotransmitter signaling: Possible non-local interactions? *International Journal of Geriatric Psychiatry*, 10, 1093–1094.
- Perry, E. K. (2002). Plants of the gods. In E. Perry, C. Ashton and A. Young (Eds.), *Neurochemistry of Consciousness* (pp. 205–224). Amsterdam: John Benjamins.
- Perry, E. K. and Perry, R. H. (1995). Acetylcholine and hallucinations: Disease-related compared with drug-induced alterations in human consciousness. *Brain and Cognition*, 28, 240–258.
- Perry, E. K., Walker, M. P., Grace, J. and Perry, R. (1999). Acetylcholine in mind: A neurotransmitter correlate of consciousness? *Trends in Neurosciences*, 22, 273–280.
- Radin, D. I. (1999). Event-related electroencephalographic correlations between isolated human subjects. *Journal of Alternative and Complementary Medicine*, 10, 315–323.
- Radin, D. I. (2008). Testing nonlocal observation as a source of intuitive knowledge. *Explore* (NY), 4, 25–35.
- Radin, D. I. (1997). *The Conscious Universe – The Scientific Truth of Psychic Phenomena*. San Francisco: HarperEdge.
- Radin, D. I. (2006). *Entangled Minds: Extrasensory experiences in a quantum reality*. New York, NY: Paraview / Simon and Shuster.
- Riba, J., Romero, S., Grasa, E., Mena, E., Carrió, I. and Barbanoj, M. J. (2006). Increased frontal and paralimbic activation following ayahuasca, the pan-Amazonian inebriant. *Psychopharmacology* (Berlin), 186, 93–98.
- Schmidt, S., Schneider, R., Utts, J. and Walach, H. (2004). Distant intentionality and the feeling of being stared at: Two meta-analyses. *British Journal of Psychology*, 95, 235–247.
- Schultes, R. E., Hofmann, A. and Ratsch, C. (2001). *Plants of the Gods*. Rochester, Vermont: Healing Arts Press.
- Schumann, T. G. (2007). Quantum physics and the dream metaphor. *NeuroQuantology*, 2, 241–252.
- Sheldrake, R. and Smart, P. (2005). Testing for telepathy in connection with e-mails. *Perceptual & Motor Skills*, 101, 771–786.
- Stuckey, D. E., Lawson, R. and Luna, L. E. (2005). EEG gamma coherence and other correlates of subjective reports during ayahuasca experiences. *Journal of Psychoactive Drugs*, 37, 163–178.
- Su, T. P., Hayashi, T. and Vaupel, D. B. (2009). When the endogenous hallucinogenic trace amine N,N-dimethyltryptamine meets the sigma-1 receptor. *Science Signal*, 2, (61): pe12.
- Voss, U., Holzmann, R., Tuin, I. and Hobson, J. A. (2009). Lucid dreaming: A state of consciousness with features of both waking and non-lucid dreaming. *Sleep*, 32, 1191–1200.
- Wackermann, J., Seiter, C., Keibel, H. and Walach, H. (2003). Correlations between brain electrical activities of two spatially separated human subjects. *Neuroscience Letters*, 336, 60–64.
- Yount, G., Smith, S., Avanzonian, V., West, J., Moore, D. and Freinkel, A. (2004). Biofield perception: A series of pilot studies with cultured human cells. *Journal of Alternative and Complementary Medicine*, 10, 463–467.
- Vollenweider, F. X. (2005). Using psilocybin to investigate the relationship between attention, working memory, and the serotonin 1A and 2A receptors. *Journal of Cognitive Neuroscience*, 17, 1497–1508.

Index

- 2-arachidonoyl-glycerol
 (2-AG) 55–56, 58
 5HT1A 311, 313
 5HT2 311, 315
- A**
- Abstract thought 76, 155, 273
 Acetylcholine 53, 66, 73, 76,
 79–80, 315, 323
 Action-perception cycle 95, 101
 Addiction XIV, 52, 57, 176,
 256–257, 310
 Affective neuroscience 120, 174,
 178, 287
 AIM model 157, 159
 Alkaloids 309, 311
 Allocentric 74–76, 201–206, 210
 Altered states of consciousness
 VII, 3, 13, 149, 286, 289–290,
 292–294, 297, 299, 315
 Alzheimer's disease 33–34, 90,
 179, 183, 185–186, 212
 Amnesia 84–85, 90, 152, 160, 291
 Amphetamine 46, 51, 243
 Amygdala 32, 39–43, 52, 54,
 58–59, 204, 207–209, 212–213,
 229, 293, 311, 313
 Anaesthesia V, XV–XVI, XXII, 3,
 12–14, 65–72, 76, 78, 95, 292
 Analgesia 14, 57, 72, 95, 262–263,
 265, 267–268
 Anandamide 55–60, 63
 Anterior cingulate gyrus 58,
 175, 194
 Anxiety 57–61, 63, 177–178, 207,
 209, 212–213, 261, 285, 298
 Anxiolytic 59
 Ascending brainstem 241
 Associative memory 96, 99
 Asymmetry XVIII, 189, 191,
 194–195, 198
- Attention XV, XVII, XX, XXIII, 10,
 24, 31, 46, 48–50, 76, 92, 94,
 96, 100, 124, 127, 130–131, 136,
 139–143, 145–147, 174–175, 180,
 182–185, 211, 216, 222, 230,
 237, 273–274, 282–285, 289,
 293–294, 311, 314–315
 Attention Deficit Hyperactivity
 disorder 284
 Attentional blink 15, 24, 28,
 94, 284
 Attractor dynamics 202–203,
 206
 Auditory 12, 135, 151, 216, 223,
 227–238, 290, 293
 Auditory hallucinations XIX, 9,
 12, 14, 133–135, 192–193, 209,
 211, 216, 223, 229, 245, 290,
 293, 299, 304
 Auditory structures 133–134
 Autoassociation 201, 203, 206,
 210
 Autobiographical 90, 122, 127,
 203, 286
 Automatic 21, 87, 94, 123, 162,
 262–263, 273
 Automatisms 92, 95
 Autonomic nervous system
 285, 291
 Awareness 17–19, 65–66, 103,
 105–109, 139–143, 151–153,
 163–164, 179–185, 241–246,
 249–250, 253–256, 286–287,
 289–292, 295, 314–316
- B**
- Barbiturate(s) 66, 76
 Basal ganglia 54, 92, 94, 266
 Behavioural XV, 14–15, 29, 52, 74,
 104, 108, 110, 170, 205, 264,
 287, 298
- Behaviourism 169–170
 Bereavement XVIII, 173–174
 Bimodality XVIII, VI, 155–159,
 163–164
 Binding 19, 31–32, 36, 87–88,
 114, 153–154, 184
 Binocular rivalry 23–24, 27, 36,
 151, 284
 Biological psychiatry 51, 171,
 178, 212–213
 Bizarreness 153–154, 246–248
 Blindness 93, 140, 142, 147–148
 Blindsight 16, 77, 93, 151
 Brainstem 45, 67, 157, 241, 244,
 292, 315
 Broadcast XVI, 87–88, 91–92, 94,
 96, 98–100, 133
 Broca's area 192
 Buddhist contemplative traditions
 282
- C**
- CA2 203–205, 208–209
 CA3 XIX, XIX, 201, 203–212
 Cannabinoid receptors
 54, 59–60, 62–63
 Cannabis 56, 58–59, 311
 Capsaicin 60
 CB1 receptor 55, 57–58, 60,
 63, 311
 CB2 54–56, 59–60, 62
 Cerebellum 10, 13, 54, 67,
 92–94
 Cerebral cortex 9, 13, 15, 20,
 27–28, 51, 54, 67, 92, 179, 191,
 212, 225, 229, 271–272, 275,
 311, 313
 Change blindness 140, 147–148
 Charles Bonnet syndrome
 VI, XIX, 215, 218, 223
 Cholecystokinin 31, 261

- Cholinergic xvi, 36, 56, 76, 150, 205, 209, 241, 244–245, 292, 311, 313, 315, 321
- Cingulate 5, 16, 23, 58, 68, 175, 194, 202–203, 218, 286, 292–294, 317
- Cocaine 49, 51
- Cochlea 228, 230, 232
- Cognition xvii–xviii, xxi, 54, 61, 92, 95, 121, 124–126, 140, 146, 158–159, 179, 247, 254–255, 263, 271, 274, 309, 314, 316
- Cognitive neuroscience xi, xi, 3, 28, 89–90, 125–127, 165, 171–172, 247, 278, 322–323
- Coherence 12, 19–20, 162, 184–185, 292, 307–308, 311, 323
- Conditioning 52, 259, 262–268
- Confabulation 114, 116, 159, 303
- Confidence decision network 117–118
- Conscious expectation 265–266, 268
- Conscious placebo responses 260
- Corpus callosum 94, 188, 194, 198
- Cortical networks 22, 50, 122, 162
- Corticothalamic 68, 70, 72, 101
- Cortisol 256–257, 264, 266, 286
- Creative xxiii, 114, 271–279, 295, 313
- Creativity vii, xx, 271–278
- Culture 126, 128, 137, 309
- Current source density (CSD) 9, 11, 14
- D**
- D1 receptors 42, 44
- D2 receptors 44
- Darwinian xx, 271–272, 275–278
- Decision criterion 103–104, 110
- Decision making xv, xvii, 50, 95, 99, 102, 120, 151
- Default mode xvii, xvii, 121, 125–127, 158–159, 165
- Default network xvii, xvii, 16, 125, 127
- Deictic 190, 192, 197
- Deixis 190
- Delirium 76, 150, 154, 159, 165, 242, 244
- Delta frequency 8
- Delusions xviii, 95, 199, 209–210, 243
- Dementia with Lewy bodies 34, 36, 179, 183, 185–186
- Dementia vi, xiv, xviii, 34, 36, 90, 95, 179–180, 182–186, 276, 278–279, 321
- Dendrites 5, 7, 12, 67
- Dentate gyrus 203–205, 207, 210
- Depression vi, xiv, xviii, 40, 48, 58–59, 61, 63, 66, 70, 169, 172–178, 186–187, 242, 246, 250, 285, 298
- Depth of processing 20, 22, 24, 105–106
- Desflurane 67
- Determinism vi, 113–114, 246
- Dissociation 90, 106–107, 143, 165, 273, 290–291, 294, 313
- Dopamine (DA) 39–49, 57, 241, 243, 246–247, 250, 262, 268
- Dopaminergic 51, 57, 61, 76, 211, 241, 243–245
- Dorsal visual stream 16, 93
- Dorsolateral prefrontal cortex 58, 163, 191, 194, 241
- Dream consciousness 153, 160, 241
- Dreaming vi–vii, xviii–xix, xxii–xxiii, 29–30, 53, 76, 78, 134, 149–165, 207, 239–242, 244–248, 314–315, 323
- Dreams 29, 133, 149, 152–154, 156, 161, 165, 207, 218, 222, 239–240, 242, 246–248, 274, 321
- E**
- Ecstasy xx, vii, 249–257, 311
- Ego xx, 163, 246, 272–273, 276–277
- Egocentric xvi, 73, 77, 202–203, 205, 272, 276–277
- Electroencephalogram (EEG) 4–5, 7, 9, 21, 23, 28–32, 68–70, 157, 160–161, 240, 246, 283–284, 292–294, 296, 303, 305–306, 314, 317
- Electrophysiological xxii, 3, 14, 31, 34, 51, 68, 227, 268, 292, 294
- Emotion(s) 40–41, 53, 113, 120, 156, 160, 178, 180, 244, 247, 287
- Emotional xix, 39–43, 74–75, 119–120, 155, 173, 178–179, 207, 229, 232, 242, 245, 260–261, 282, 286, 291, 311, 314, 321
- Endocannabinoid xv, v, 53–54, 56–63
- Endocrine system 264, 285
- Endorphins 57, 176
- Entanglement xxi, 306–308, 319
- Entorhinal cortex 40, 52, 75–76, 79–80, 202–205, 207
- Epilepsy 188, 290, 294
- Episodic buffer 181
- Episodic memory 25, 41, 73–74, 76, 79–80, 85, 96, 203, 205, 207
- Ergoline 311
- Etomidate 66–67
- Euphoria 57, 311
- Event related potentials 130
- Evolutionary 10, 13, 54, 101, 119, 153–154, 158, 176, 188, 195, 198, 234, 248
- Expectation 46, 259, 261–262, 264–268, 305, 317
- Explicit v, xvi, 83–90, 103–106, 110, 114–116, 122–123, 127, 138, 196, 206, 249, 254, 283, 286
- Externality label 228, 233–236
- Extra-sensory perception 306, 317
- F**
- False memories 153
- Familiarity 86, 88, 209, 239
- fMRI 3–5, 7–8, 10–15, 86, 104, 110, 124, 148, 162, 209, 212, 224, 278, 293, 298, 306, 308, 320
- Free will vi, xvii, 113, 115–116, 119
- Frontal xix, xx, 16, 19, 22–24, 27–28, 32, 41, 54, 69, 71, 84, 120, 127, 161–162, 185, 187, 218, 220, 229, 241–242, 276–277, 292–293, 296, 311, 313–315, 317, 323

- Frontoparietal 67, 71, 94
 Frontotemporal dementia
 185, 276, 278–279
 Functional magnetic resonance
 imaging (fmri) 3, 130, 212–
 213, 225, 238, 293, 308, 317, 322
 Fusiform facial area 69, 216
- G**
 GABA 31, 33–34, 53–54, 66–68,
 70, 208–209, 211, 238, 245, 313
 GABAergic 34, 41, 45, 76–77, 201,
 205, 208–209, 211, 245, 292
 Gamma oscillations v, xv, xix,
 15, 20, 22–24, 29–37, 68–71,
 77–78, 94, 100, 184, 205, 210,
 230, 232–233, 237, 311, 314
 Ganzfeld 303–304, 307–308
 Gaze 123–124, 127, 142–143, 145,
 147–148
 Gestalt 210
 Global Workspace Theory
 xvi, 91–92, 96, 99, 101, 133,
 218, 220
 Glutamate 37, 40, 45, 53–54, 60,
 66, 208–209, 294, 311, 313
 Glutamic acid decarboxylase
 (GAD) 33, 208
 Growth hormone 264, 266
- H**
 Hallucination(s) xiv, xviii–xix,
 32, 36, 80, 95, 133, 154, 159,
 179, 183, 185, 192–193, 201,
 207–213, 215–225, 227,
 229–230, 232–233, 235–236,
 238, 243–245, 304, 309, 311,
 313, 315, 322–323
 Hallucinatory experiences
 149, 236, 242, 315
 Hallucinatory phenomena
 225, 239, 244, 247
 Hallucinogens 315
 Hard problem 26, 152
 Heroin 57
 Higher order syntactic thought
 (HOST) 114, 117, 119, 173
 Hippocampal xix, 36–37, 41,
 43–44, 47–49, 51–52, 79,
 85–87, 89, 98, 173, 202, 204–
 205, 208–213, 294
 Hippocampus xv, xvi, 25, 28,
 31–32, 35–37, 39–40, 42–43,
 45–46, 48, 51–52, 54, 58,
 67, 75, 79, 83–90, 201–202,
 204–205, 207–213, 229, 244,
 293, 311, 313
 Histaminergic 292
 Hypnosis 164, 256, 290, 294, 296
 Hypoglycaemia 264
- I**
 Illusions vi, 139, 141, 144–145,
 147, 151, 227, 238
 Imagery xix, 201, 203, 206,
 212, 218, 228–229, 233–237,
 240, 311
 Imaging x–xi, xiii, xxi, 3, 7, 10,
 19, 27, 46, 51, 84, 130, 148, 160,
 163–164, 175, 212–213, 215–217,
 224–225, 237–238, 243, 262,
 268, 285, 292–294, 308,
 310–311, 317, 320, 322
 Immune system 263, 285
 Implicit v, xvi–xvii, 83–90,
 103–107, 110, 113–116, 122–123,
 127, 136, 138, 196, 245
 Inattention blindness
 142, 147–148
 Information integration theory
 70
 Intermediate level theory 134
 Internal representation
 xvi, 73–76, 121
 Interneurons xix, 7, 31, 33–37,
 51, 201, 208–209, 211–212
 Isoflurane 67, 79
- J**
 Judgement 182, 274
- K**
 Ketamine 67, 294
 Kundalini 286, 293
- L**
 Language vi, xvii–xix, 78, 94,
 115, 117, 122, 129–135, 137–138,
 140, 187–192, 194–199, 207,
 223, 228, 271, 273, 314
 Learning xvii, 39, 43, 54, 74,
 79–80, 84, 86, 89–90, 101,
 110, 120, 130–131, 136–138, 170,
 205, 211–213, 262–263, 268,
 277, 303
 LIDA 91–92, 95, 97–101
 Limbic xv, 39–40, 45, 49, 51,
 54, 101, 194, 223, 238, 241,
 243–244, 293, 295, 311, 313
 Linguistic sign 189, 196
 Linguistic theory 130
 Local field potentials (lfps)
 3, 205
 Long term memory 24, 75, 78,
 83, 132, 181, 244
 Long-term working memory
 95, 101
 Long-term depression (LTD)
 48–49, 198
 Long-term potentiation (LTP)
 47, 203, 208, 211
 LSD 249, 290, 294, 311, 314,
 316, 322
 Lucid dreaming vi, xviii–xix,
 155, 159–165, 239, 246, 248,
 314, 323
- M**
 Magic xxii, 139–142, 144–148,
 273, 311
 Magnetic resonance imaging
 3, 130, 212–213, 225, 238, 285,
 293, 308, 317, 322
 Manic depressive disorder 187
 Masking 21, 24–25, 94, 106–107,
 120
 MDMA vii, xx, 249–257
 Medial geniculate body 230
 Medial temporal lobe 25, 84,
 86–87, 89, 194, 202, 209, 212
 Meditation vii, xx–xxi, 281–288,
 290–294, 296, 298–299, 313
 Memory v, xvi, xix, 22, 24–25,
 27–28, 31, 33–34, 39, 41–42, 51,
 54, 59, 73–80, 83–90, 94–96,
 98–99, 101, 110, 114, 117,
 119–120, 122, 127, 132–133, 152,
 156, 159–160, 180–186, 195,
 201, 203–208, 212–213, 218,
 221–222, 224, 228–229, 244,
 250, 252–253, 255, 257, 291, 313,
 322–323
 Mesolimbic pathway 57

- Meta-memory 117
- Midbrain 44, 50, 175, 293
- Mindfulness based stress reduction 285
- Mirror neuron 125
- Misdirection 140–143, 147–148
- Modular Online Growth and Use of Language (MOGUL) 129, 131, 133, 136
- Monoamines 53, 55
- Morphine 57, 311
- Morphosyntax 131–132
- Motivation 40–41, 49, 54, 101, 124–125, 173, 251
- Multiple personality disorder 164, 291
- Mu-opioid 175
- Muscarinic cholinergic receptors 76
- Musical hallucinations 229, 232–233, 235
- Mystical experiences 314
- N**
- Neglect 77, 93, 171, 179
- Network oscillations 29, 34, 36–37, 75
- Neural correlates of consciousness XIII, 15, 27–28, 37, 65, 71, 164, 185, 217, 224, 227–228, 237
- Neuroimaging IX, 3, 8, 68, 124, 126, 207, 209, 224–225, 227, 235, 247, 256, 287, 292–293, 298, 322
- Neuronal assemblies 17, 19, 23, 205
- Neuronal synchrony 17, 19, 21, 24, 26, 28
- Neuroplasticity XX, 37, 281, 283
- Never-conscious XVI, 91–92, 100
- Nitrous oxide 67–68
- NMDA 37, 45, 50–51, 208–209, 294, 311
- Nocebo 261, 265, 267–268
- Noradrenaline 53–54, 250, 294
- Noradrenergic 61, 76, 241, 292
- Norepinephrine 172
- NREM sleep 240, 248
- Nucleus accumbens XV, 39–45, 51–52, 57, 245, 262, 313
- O**
- Occipital 22–23, 68, 191, 220–221
- OM meditation 284, 286, 290, 293
- Operating characteristics VI, 103–105, 107, 110
- Opioid receptors 54, 57, 60, 63, 175, 262, 313
- Opsin molecules 35
- Optimistic beliefs XX, VII, 249–250
- Optogenetic 29, 31, 34–36
- Orbitofrontal 204
- Oscillation(s) V, XV, XIX, 14, 20, 22–24, 26–37, 71, 75, 77, 79, 93–94, 184, 205–207, 211–212, 230, 232–233, 237, 243, 318
- Oxytocin 175
- P**
- Pain 53–54, 56–63, 170, 172, 174–176, 178, 195, 259, 261–262, 265–266, 268, 282, 285–287, 307
- Papez 202–203, 205
- Parahippocampal place area 216
- Parietal 9, 11, 22–24, 68, 77, 93, 125, 161–163, 202–206, 218, 229, 283, 293, 313
- Parkinson's disease 39, 58, 262, 266, 268
- Parvalbumin 31, 33–37, 51, 205, 208, 213
- Pedunculopontine tegmentum (pptg) 43, 45, 47
- Perception XVII–XIX, XXI, 10–11, 14, 16, 20, 23–25, 27–29, 31–33, 36–37, 54, 59, 65–66, 73, 77, 92, 95, 101, 103–106, 108–110, 120, 124, 133, 135, 139–141, 144–148, 152–153, 156, 160, 180, 182, 185, 187, 190, 196, 201, 206–207, 213, 225, 227–230, 232–233, 235–238, 286, 289, 291, 301, 306–307, 311, 313–318, 322–323
- Perirhinal 86, 207
- PET studies 241
- Phase 8, 10, 20, 22–23, 28, 32, 70, 75, 77–79, 86, 174, 197, 205, 212, 240, 251, 291
- Phase locking 23, 32
- Phonological engrams 190, 192–193
- Phonological loop 183
- Phonological store 132
- Phonology 131, 192
- Place cells 40, 52, 212
- Placebo VII, XIV, XX, XXIII, 259–268, 314, 322
- Planning 73–76, 114, 116, 151, 192, 229, 253
- Post-decision wager(ing) XVII, 103, 108–111
- Preconscious XVI, 14, 27, 36, 53, 91–92, 96–101, 151
- Precuneus 67–68, 163, 202–203, 292, 298, 317
- Prefrontal cortex XV, XIX, 12, 19, 28, 34, 39–44, 47–48, 58, 76, 79, 94, 124, 128, 134, 163, 165, 191, 194, 210, 241, 244, 292–293
- Primary consciousness XVIII, 155–156, 158–159, 162, 164, 195
- Priming 22, 27, 83–85, 87–90, 136
- Procedural memory 84, 98–99
- Propofol 12, 66–67, 69–71
- Prosopagnosia 93
- Psi 307–308, 314, 316–320, 322
- Psilocybine 311, 314, 321
- Psychopathology XIX, 199, 211, 242–243, 247
- Psychosis VI–VII, XIV, XIX, 51, 58, 187–188, 190, 195–196, 198, 201, 211–213, 225, 233, 239, 241–248, 310
- Psychosocial context 259–260
- Psychotherapy 253, 259–260, 267
- Psychotic 150, 243
- Psychotic symptoms 188, 190, 194–195, 243, 246, 315
- Pyramidal cells 5, 7, 31–32, 35, 194, 204–205, 208–211
- Pyramidal neurons 3, 8, 203–205, 207–208, 210
- Q**
- Qualia XIII, 114, 318

- Quantum entanglement 306, 319
- Quantum physics XIII, XXI, 301, 318–319, 323
- R**
- Recall 37, 41, 83–84, 86, 114, 203–206, 212, 235, 240
- Recognition 28, 83–84, 86, 88–89, 96, 164, 204, 206, 212, 234, 247, 253–254
- Recollection XVI, 83, 86–89, 203, 212
- Reflective consciousness XVIII, XXII, 155, 215
- Reflexive 123, 147
- Reinforcement learning theory 74
- Relaxation 282, 290, 294, 298–299
- REM sleep 16, 76, 150, 157–165, 205, 207, 240–241, 245, 248, 315
- Retinocentric 202–203
- Reward XVII, 40, 44–46, 48–49, 52, 54, 57, 61, 63, 113–114, 117, 127, 178, 204–205, 213, 241, 261–262, 268, 311
- S**
- Schizophrenia VI, XIV, XVIII–XIX, 29, 32–34, 37, 39–40, 46, 50–51, 164, 187–189, 193–194, 197–199, 201, 207–213, 220, 223, 225, 229, 233, 238, 242–248
- Second language XVII, 130–131, 137
- Secondary consciousness XX, 155–156, 158–160, 163–164
- Secondary representational system 195–196
- Self awareness 197
- Self consciousness XXII
- Self-induced altered states of consciousness VII, 289
- Self induced hypnosis 294
- Semantic memory 86, 90
- Sensory XVI, 7, 10, 16, 18–19, 31–32, 36, 42, 54, 56, 59–60, 67, 69, 74–76, 78, 92–94, 96, 98–99, 103–105, 108–109, 111, 114, 133–134, 140, 143–144, 149–150, 191–193, 196, 202–205, 207, 209, 211, 218–219, 223, 241–242, 244–246, 260, 285, 290–291, 295, 306–308, 313, 316–317, 320
- Sensory deprivation 246, 290
- Separation distress 174–175, 177–178
- Serotonergic 58, 61, 76, 241, 256, 292, 313, 315, 321–322
- Serotonin XVIII, 50, 53–54, 59, 172–173, 177, 250, 264, 294, 313–315, 322–323
- Serotonin receptors 54, 314
- Shamanic VII, XXI, 309–311, 315–316, 318, 320–321
- Sigma receptors 313
- Signal detection theory 104, 109–111
- Sleep XVIII–XIX, XXV, 12–13, 16, 29–30, 32, 53–54, 61, 67, 71, 76, 78, 93–95, 102, 149–151, 153, 155–165, 172, 184, 205, 207, 212, 215, 218, 239–242, 244–248, 250–252, 255–256, 289–290, 292, 295–296, 298, 315, 323
- Slow cortical potential (SCP) V, XV, 3–4, 13–14
- Slow wave sleep 30, 76, 292, 298
- Small world networks 21
- Social consciousness VI, XVII, 121–123, 125–126
- Social context 123, 127, 260
- Social gaze 123, 127
- Social loss 174, 176–178
- Social phobia 207, 213
- Split brain 163, 165, 188, 197
- SSRIs 59, 173
- Stellate cells 5, 7, 79
- Striatum XIX, 39–41, 45, 124, 311, 313
- Subiculum 39–49, 52, 202
- Subjective experience 121–122, 170, 223, 228, 239–241, 281, 283, 290–291, 307, 310, 318
- Subjectivity XX, 153–154, 211, 241, 247, 281, 286–287
- Subliminal 14, 22, 27, 36, 92, 94, 101, 135
- Substantia nigra 44, 266
- Suicide 58, 63, 170, 172
- Synchronisation 30, 32, 36, 210
- Synchronous 30–32, 68, 78, 205
- Synchrony 17, 19–21, 23–24, 26–28, 32–33, 35, 37, 68, 230, 238, 287, 314
- Syntactic structure(s) 132
- T**
- Telepathy XXI, XXIII, 301–304, 306–308, 316–317, 323
- Tetrahydrocannabinol (THC) 56–58, 312
- Thalamic inputs 5–7, 10
- Thalamocortical 7, 9, 91, 158, 211, 238, 292
- Thalamus 7, 42, 54, 59, 67–68, 70, 229, 266, 292, 311, 313, 315
- Thalamic 5–7, 10, 12–15, 40, 42–43, 67–68, 71–72, 175, 202, 211–212, 268, 293
- Theory of mind 124, 127, 185
- Theta 15, 22–24, 27, 37, 47, 75, 79, 94, 98, 100, 212, 292, 294
- Threat simulation theory 153
- Time V, XIV, 71–78, 121, 190–197, 201–209, 301, 307
- Tinnitus XIX, 32, 36, 227, 229, 232–233, 235–236, 238
- Torque XVIII–XIX, 187, 189–191, 194
- Torrence tests 274
- Transcendental meditation 290–292, 296, 298–299
- U**
- Unconscious V–VII, XV–XVII, XIX–XX, XXII, 8, 12, 15–16, 19–21, 24–28, 53–54, 56, 61, 65, 78, 88, 91–94, 98–100, 103–107, 109–110, 113–114, 120, 129, 133–137, 151, 156, 210, 215–216, 219, 221–224, 227–229, 234–235, 237, 240, 244, 249–250, 254–255, 259–260, 262–263, 265–266, 268, 276, 301, 305–307, 320
- Unconscious placebo responses 263, 265

V

Vanilloid receptors 60
 Vegetative state xxiii, 12, 14, 66, 298
 Ventral striatum xix, 40–41, 45, 124
 Ventral tegmental area 43–44, 211, 313
 Ventral visual stream 207
 Verbal hallucinations xix, 223, 233, 235–236
 Vipassana 282, 284
 Vision 15, 105, 110, 119, 145, 215, 217–218, 220, 222, 224–225, 230, 242

Visions 224, 309, 313–316, 318, 320
 Visual awareness xvii, 110, 139–140, 143, 147, 151–153
 Visual cortex 6, 27, 36, 71, 77, 93, 104–105, 111, 151, 204, 218, 223, 306, 311, 313
 Visual hallucinations xix, 159, 185, 209, 211–212, 215–216, 218–219, 221–225, 244, 322
 Volition xviii, 11, 14, 99, 155–156, 163–164

W

Waking 78, 95, 150, 153–165, 180, 188, 206, 212, 240, 247–248, 289, 311, 315, 323
 Wernicke's area 130, 189, 192
 Working memory 22, 27, 33–34, 39, 42, 75, 77–79, 94–95, 101, 132–133, 181–185, 322–323

Y

Yoga 291, 293, 295, 297–298

Z

Zen 285, 288, 296, 298

Advances in Consciousness Research

A complete list of titles in this series can be found on the publishers' website, www.benjamins.com

- 80 MÖLDER, Bruno: Mind Ascribed. An elaboration and defence of interpretivism. 2010. xii, 293 pp.
- 79 PERRY, Elaine, Daniel COLLERTON, Fiona E.N. LEBEAU and Heather ASHTON (eds.): New Horizons in the Neuroscience of Consciousness. 2010. xxv, 330 pp.
- 78 CZIGLER, István and István WINKLER (eds.): Unconscious Memory Representations in Perception. Processes and mechanisms in the brain. 2010. x, 274 pp.
- 77 GLOBUS, Gordon G.: The Transparent Becoming of World. A crossing between process philosophy and quantum neurophilosophy. 2009. xiii, 169 pp.
- 76 BRÄTEN, Stein: The Intersubjective Mirror in Infant Learning and Evolution of Speech. 2009. xxii, 351 pp.
- 75 SKRBINA, David (ed.): Mind that Abides. Panpsychism in the new millennium. 2009. xiv, 401 pp.
- 74 CAÑAMERO, Lola and Ruth AYLETT (eds.): Animating Expressive Characters for Social Interaction. 2008. xxiii, 296 pp.
- 73 HARDCASTLE, Valerie Gray: Constructing the Self. 2008. xi, 186 pp.
- 72 JANZEN, Greg: The Reflexive Nature of Consciousness. 2008. vii, 186 pp.
- 71 KROIS, John Michael, Mats ROSENGREN, Angela STEIDELE and Dirk WESTERKAMP (eds.): Embodiment in Cognition and Culture. 2007. xxii, 304 pp.
- 70 RAKOVER, Sam S.: To Understand a Cat. Methodology and philosophy. 2007. xviii, 253 pp.
- 69 KUCZYNSKI, John-Michael: Conceptual Atomism and the Computational Theory of Mind. A defense of content-internalism and semantic externalism. 2007. x, 524 pp.
- 68 BRÄTEN, Stein (ed.): On Being Moved. From mirror neurons to empathy. 2007. x, 333 pp.
- 67 ALBERTAZZI, Liliana (ed.): Visual Thought. The depictive space of perception. 2006. xii, 380 pp.
- 66 VECCHI, Tomaso and Gabriella BOTTINI (eds.): Imagery and Spatial Cognition. Methods, models and cognitive assessment. 2006. xiv, 436 pp.
- 65 SHAUMYAN, Sebastian: Signs, Mind, and Reality. A theory of language as the folk model of the world. 2006. xxvii, 315 pp.
- 64 HURLBURT, Russell T. and Christopher L. HEAVEY: Exploring Inner Experience. The descriptive experience sampling method. 2006. xii, 276 pp.
- 63 BARTSCH, Renate: Memory and Understanding. Concept formation in Proust's *A la recherche du temps perdu*. 2005. x, 160 pp.
- 62 DE PREESTER, Helena and Veroniek KNOCKAERT (eds.): Body Image and Body Schema. Interdisciplinary perspectives on the body. 2005. x, 346 pp.
- 61 ELLIS, Ralph D.: Curious Emotions. Roots of consciousness and personality in motivated action. 2005. viii, 240 pp.
- 60 DIETRICH, Eric and Valerie Gray HARDCASTLE: Sisyphus's Boulder. Consciousness and the limits of the knowable. 2005. xii, 136 pp.
- 59 ZAHAVI, Dan, Thor GRÜNBAUM and Josef PARNAS (eds.): The Structure and Development of Self-Consciousness. Interdisciplinary perspectives. 2004. xiv, 162 pp.
- 58 GLOBUS, Gordon G., Karl H. PRIBRAM and Giuseppe VITIELLO (eds.): Brain and Being. At the boundary between science, philosophy, language and arts. 2004. xii, 350 pp.
- 57 WILDGEN, Wolfgang: The Evolution of Human Language. Scenarios, principles, and cultural dynamics. 2004. xii, 240 pp.
- 56 GENNARO, Rocco J. (ed.): Higher-Order Theories of Consciousness. An Anthology. 2004. xii, 371 pp.
- 55 PERUZZI, Alberto (ed.): Mind and Causality. 2004. xiv, 235 pp.
- 54 BEAUREGARD, Mario (ed.): Consciousness, Emotional Self-Regulation and the Brain. 2004. xii, 294 pp.
- 53 HATWELL, Yvette, Arlette STREIER and Edouard GENTAZ (eds.): Touching for Knowing. Cognitive psychology of haptic manual perception. 2003. x, 322 pp.
- 52 NORTHOFF, Georg: Philosophy of the Brain. The brain problem. 2004. x, 433 pp.
- 51 DROEGE, Paula: Caging the Beast. A theory of sensory consciousness. 2003. x, 183 pp.
- 50 GLOBUS, Gordon G.: Quantum Closures and Disclosures. Thinking-together postphenomenology and quantum brain dynamics. 2003. xxii, 200 pp.
- 49 OSAKA, Naoyuki (ed.): Neural Basis of Consciousness. 2003. viii, 227 pp.
- 48 JIMÉNEZ, Luis (ed.): Attention and Implicit Learning. 2003. x, 385 pp.
- 47 COOK, Norman D.: Tone of Voice and Mind. The connections between intonation, emotion, cognition and consciousness. 2002. x, 293 pp.

- 46 MATEAS, Michael and Phoebe SENGERS (eds.): Narrative Intelligence. 2003. viii, 342 pp.
- 45 DOKIC, Jérôme and Joëlle PROUST (eds.): Simulation and Knowledge of Action. 2002. xxii, 271 pp.
- 44 MOORE, Simon C. and Mike OAKSFORD (eds.): Emotional Cognition. From brain to behaviour. 2002. vi, 350 pp.
- 43 DEPRAZ, Nathalie, Francisco J. VARELA and Pierre VERMERSCH: On Becoming Aware. A pragmatics of experiencing. 2003. viii, 283 pp.
- 42 STAMENOV, Maxim I. and Vittorio GALLESE (eds.): Mirror Neurons and the Evolution of Brain and Language. 2002. viii, 392 pp.
- 41 ALBERTAZZI, Liliana (ed.): Unfolding Perceptual Continua. 2002. vi, 296 pp.
- 40 MANDLER, George: Consciousness Recovered. Psychological functions and origins of conscious thought. 2002. xii, 142 pp.
- 39 BARTSCH, Renate: Consciousness Emerging. The dynamics of perception, imagination, action, memory, thought, and language. 2002. x, 258 pp.
- 38 SALZARULO, Piero and Gianluca FICCA (eds.): Awakening and Sleep–Wake Cycle Across Development. 2002. vi, 283 pp.
- 37 PYLKKÄNEN, Paavo and Tere VADÉN (eds.): Dimensions of Conscious Experience. 2001. xiv, 209 pp.
- 36 PERRY, Elaine, Heather ASHTON and Allan H. YOUNG (eds.): Neurochemistry of Consciousness. Neurotransmitters in mind. With a foreword by Susan Greenfield. 2002. xii, 344 pp.
- 35 Mc KEVITT, Paul, Seán Ó NUALLÁIN and Conn MULVIHILL (eds.): Language, Vision and Music. Selected papers from the 8th International Workshop on the Cognitive Science of Natural Language Processing, Galway, 1999. 2002. xii, 433 pp.
- 34 FETZER, James H. (ed.): Consciousness Evolving. 2002. xx, 253 pp.
- 33 YASUE, Kunio, Mari JIBU and Tarcisio DELLA SENTA (eds.): No Matter, Never Mind. Proceedings of Toward a Science of Consciousness: Fundamental approaches, Tokyo 1999. 2002. xvi, 391 pp.
- 32 VITIELLO, Giuseppe: My Double Unveiled. The dissipative quantum model of brain. 2001. xvi, 163 pp.
- 31 RAKOVER, Sam S. and Baruch CAHLON: Face Recognition. Cognitive and computational processes. 2001. x, 306 pp.
- 30 BROOK, Andrew and Richard C. DEVIDI (eds.): Self-Reference and Self-Awareness. 2001. viii, 277 pp.
- 29 VAN LOOCKE, Philip (ed.): The Physical Nature of Consciousness. 2001. viii, 321 pp.
- 28 ZACHAR, Peter: Psychological Concepts and Biological Psychiatry. A philosophical analysis. 2000. xx, 342 pp.
- 27 GILLET, Grant R. and John McMILLAN: Consciousness and Intentionality. 2001. x, 265 pp.
- 26 Ó NUALLÁIN, Seán (ed.): Spatial Cognition. Foundations and applications. 2000. xvi, 366 pp.
- 25 BACHMANN, Talis: Microgenetic Approach to the Conscious Mind. 2000. xiv, 300 pp.
- 24 ROVEE-COLLIER, Carolyn, Harlene HAYNE and Michael COLOMBO: The Development of Implicit and Explicit Memory. 2000. x, 324 pp.
- 23 ZAHAVI, Dan (ed.): Exploring the Self. Philosophical and psychopathological perspectives on self-experience. 2000. viii, 301 pp.
- 22 ROSSETTI, Yves and Antti REVONSUO (eds.): Beyond Dissociation. Interaction between dissociated implicit and explicit processing. 2000. x, 372 pp.
- 21 HUTTO, Daniel D.: Beyond Physicalism. 2000. xvi, 306 pp.
- 20 KUNZENDORF, Robert G. and Benjamin WALLACE (eds.): Individual Differences in Conscious Experience. 2000. xii, 412 pp.
- 19 DAUTENHAHN, Kerstin (ed.): Human Cognition and Social Agent Technology. 2000. xxiv, 448 pp.
- 18 PALMER, Gary B. and Debra J. OCCHI (eds.): Languages of Sentiment. Cultural constructions of emotional substrates. 1999. vi, 272 pp.
- 17 HUTTO, Daniel D.: The Presence of Mind. 1999. xiv, 252 pp.
- 16 ELLIS, Ralph D. and Natika NEWTON (eds.): The Caldron of Consciousness. Motivation, affect and self-organization — An anthology. 2000. xxii, 276 pp.
- 15 CHALLIS, Bradford H. and Boris M. VELICHKOVSKY (eds.): Stratification in Cognition and Consciousness. 1999. viii, 293 pp.
- 14 SHEETS-JOHNSTONE, Maxine: The Primacy of Movement. 1999. xxxiv, 583 pp.
- 13 VELMANS, Max (ed.): Investigating Phenomenal Consciousness. New methodologies and maps. 2000. xii, 381 pp.

- 12 **STAMENOV, Maxim I. (ed.):** Language Structure, Discourse and the Access to Consciousness. 1997. xii, 364 pp.
- 11 **PYLKKÖ, Pauli:** The Aconceptual Mind. Heideggerian themes in holistic naturalism. 1998. xxvi, 297 pp.
- 10 **NEWTON, Natika:** Foundations of Understanding. 1996. x, 211 pp.
- 9 **Ó NUALLÁIN, Seán, Paul Mc KEVITT and Eoghan Mac AOGÁIN (eds.):** Two Sciences of Mind. Readings in cognitive science and consciousness. 1997. xii, 490 pp.
- 8 **GROSSENBACHER, Peter G. (ed.):** Finding Consciousness in the Brain. A neurocognitive approach. 2001. xvi, 326 pp.
- 7 **MAC CORMAC, Earl and Maxim I. STAMENOV (eds.):** Fractals of Brain, Fractals of Mind. In search of a symmetry bond. 1996. x, 359 pp.
- 6 **GENNARO, Rocco J.:** Consciousness and Self-Consciousness. A defense of the higher-order thought theory of consciousness. 1996. x, 220 pp.
- 5 **STUBENBERG, Leopold:** Consciousness and Qualia. 1998. x, 368 pp.
- 4 **HARDCASTLE, Valerie Gray:** Locating Consciousness. 1995. xviii, 266 pp.
- 3 **JIBU, Mari and Kunio YASUE:** Quantum Brain Dynamics and Consciousness. An introduction. 1995. xvi, 244 pp.
- 2 **ELLIS, Ralph D.:** Questioning Consciousness. The interplay of imagery, cognition, and emotion in the human brain. 1995. viii, 262 pp.
- 1 **GLOBUS, Gordon G.:** The Postmodern Brain. 1995. xii, 188 pp.