

Introduction. Mental processes in the human brain

For centuries, the relation of the human mind to the brain has been debated. How can seemingly immaterial entities such as thoughts and memories arise from biological material? Advances in neuroscience have now led to wide acceptance in science and medicine that *all* aspects of our mental life—our perceptions, thoughts, memories, actions, plans, language, understanding of others and so on—in fact depend upon brain function.

In addition to being beneficiaries of the brain's complex functioning, people can also be victims of this. Many devastating and disabling conditions are a consequence of disrupted brain function, as in cases of dementia or following a stroke. Specific cognitive functions can be severely impaired, even while others remain intact in the same person. Disrupted brain function is also increasingly thought to underlie the major mental illnesses. Studies of human brain function (together with related animal studies) are thus critical for understanding major neurological and psychiatric disease. Hence, this field has become a key part of biomedical science.

In addition to the biomedical approach, studies of the human mind and brain have also benefited greatly from psychological approaches. These originally grew out of philosophy of mind, but then became determinedly experimental. More recently, a further key approach has involved computational modelling of cognitive functions in the brain. This approach has some historical roots in the development of intelligent machines during the computer revolution, but has since become a sophisticated mathematical branch of neuroscience. Nowadays, most cutting-edge research on human brain function fuses the three very different traditions or strands together (i.e. biomedical, psychological and computational), in a highly interdisciplinary field. Scientific study of the human mind and brain has apparently come of age in the past decade or so, with a series of remarkable methodological breakthroughs, and theoretical advances, in addition to an ever-growing number of empirical findings.

Space constraints here preclude a comprehensive review of how the current layout of the field has arisen for study of mental processes in the human brain. Nevertheless, several historical markers can be identified approximately. The computer revolution of the 1940s led in turn to a 'cognitive revolution' in psychology during the 1950s and 1960s, with the focus upon information processing (via analogies to computers and programs) leading to an interest in internal mental processes, rather than just in the overt behaviour that had been the dominant concern of the preceding 50 years.

One contribution of 14 to a Discussion Meeting Issue 'Mental processes in the human brain'.

While studies of lower-level sensory and motor processes have been fairly well integrated with underlying physiology for over a century, this was not always so for higher mental processes. A student in the mid-twentieth century might have been taught simply that 'association cortex' is involved in higher mental processes, in some non-specific (or 'mass action') way. This view often prevailed back then, even though Broca & Wernicke had reported on rather specific language deficits after particular brain damage in neurological patients considerably earlier (late nineteenth century). Several key developments were to bring the neuroscience of higher mental processes into focus again, with a particular emphasis on specificity in the underlying brain mechanisms.

One development was that advances from cognitive psychology, using its information-processing framework, led to new insights into the selective deficits of brain-damaged patients. The highly selective form of amnesia observed by [Scoville & Milner \(1957\)](#), after bilateral temporal lobe surgery in patient HM, provided one particularly striking example of specificity. Information-processing models from cognitive psychology were then used to provide further insights into highly selective cognitive deficits in a variety of domains, including not only long-term memory but also short-term memory, semantic memory, reading, planning and so on. This led to the new field of cognitive neuropsychology in the 1970s and 1980s (see [McCarthy & Warrington 1988](#); [Shallice 1988](#), for reviews).

In an overlapping period, an independent but equally critical development was that single-cell recording methods for studying neural activity in animals, which had originally been applied during anaesthesia (e.g. [Hubel & Wiesel 1959](#)), began to be used in awake behaving animals as they performed increasingly complex tasks. It became possible to relate response properties of neurons to more 'cognitive' issues, such as coding the particular place that an exploring animal was currently located in (e.g. [O'Keefe & Dostrovsky 1971](#)); perceptual discrimination ([Newsome & Britten 1989](#)); or even perceptual awareness ([Logothetis & Schall 1989](#)), as opposed to purely stimulus-driven responses; selective attention ([Moran & Desimone 1985](#)); working memory ([Fuster *et al.* 1985](#)) and so on.

As regards computational modelling, connectionist models of cognitive functions emerged in the 1980s. These sought to incorporate elementary aspects of cellular assemblies, using a so-called 'brain analogy', rather than the longstanding and rather literal computer analogy used hitherto by many information-processing approaches (e.g. [McClelland & Rumelhart 1985](#)). Connectionist models were also often strongly influenced by findings and topics from cognitive psychology and neuropsychology (e.g. [Hinton & Shallice 1991](#)).

More recent computational theories now incorporate increasing cellular and neurotransmitter detail (e.g. Dayan & Abbot 2005; see also Cohen *et al.* 2007). Indeed, it is arguably only since the 1990s that the biomedical, psychological and computational strands have become very closely interwoven. Prior to then, the methods of the time rarely allowed localization of function to be studied with high resolution in brain-damaged patients, while original connectionist models typically bore only a rather abstract similarity to actual neural populations.

A critical further development that has led to substantial advances, particularly for studies of the human brain, was the advent of new methods for non-invasive measurement of activity within the human brain. A series of technical breakthroughs led to increasingly widespread use of positron emission tomography (PET) in the 1980s and subsequently to functional magnetic resonance imaging (fMRI) from the 1990s. In addition to technological advances with such methods, a further key aspect was their application to human volunteers engaged in different cognitive tasks drawn from experimental psychology (Posner & Raichle 1994). Indeed, while there have since been many mathematical advances in the techniques used for analysing neuroimaging data (e.g. Valdes-Sosa *et al.* 2005), the combination of neural measures with psychological methods has remained critical. Even the most technically sophisticated neuroimaging approaches may be of little use for studying cognition, unless applied to carefully chosen paradigms designed to highlight one or another aspect of cognition, and to fractionate this into component processes. Methods from cognitive psychology and psychophysics (and, more recently, even from economics) have thus contributed much to recent advances in neuroimaging of human cognitive function, just as they have been critical for neuropsychology, in an increasingly interdisciplinary field.

The advent of PET and fMRI triggered an explosion of interest in relating cognitive function to human brain activity. This also rekindled interest in some existing methods that can provide greater temporal resolution, such as electroencephalography (EEG), and related but technically more complex methods such as magnetoencephalography (MEG). At around the same time, separate developments in reductionist neuroscience studies at the molecular level, in relatively simple animals, were also being related to cognitive function (such as memory), with some spectacular successes (e.g. Kandel 2004). Molecular variations at the genetic level are now being related even to neural activity across the whole brain, in human neuroimaging (Hariri *et al.* 2006). Thus, there is an ever-increasing tendency for neuroscience studies at a variety of different levels to be related to each other, with all levels being linked to cognitive function. The study of mental processes in the human brain is now based on a convergence of scientific traditions, together with enabling methods and new technologies.

The interdisciplinarity of the current field is further illustrated by the growing importance of formal mathematical models for cognitive functions, which have evolved from the connectionist networks of the

1980s through to more detailed theoretical approaches that integrate data from cellular and neurotransmitter levels also (Dayan & Abbot 2005). Such formal models are increasingly being used to derive explicit predictions for neuroimaging studies, a development that we strongly welcome, as exemplified by several contributions in the present volume (e.g. Cohen *et al.* 2007; Dolan 2007; Kouider & Dehaene 2007). Studies of specific cognitive deficits in patients with selective brain damage still continue to provide essential information (e.g. Burgess *et al.* 2007; D'Esposito 2007; Patterson 2007; Robbins 2007; Stuss & Alexander 2007; Vuilleumier & Driver 2007), which can fruitfully be related to computational models of cognitive function and to neural networks. More recently, studies of brain-damaged patients can also include functional neuroimaging in the patients themselves, to assess the impact of their focal lesions upon function in remote but interconnected regions that survive the lesion (e.g. D'Esposito 2007; Vuilleumier & Driver 2007). This provides a new approach for understanding network interactions between communicating brain areas.

A further methodological innovation involves the use of transcranial magnetic stimulation (TMS; Walsh & Pascual-Leone 2003), as a means for non-invasive stimulation of particular brain regions, which can have highly selective (and transient) effects on normal cognitive function. This method allows causal manipulation of activity in particular brain regions, offering perhaps the first such method for humans (albeit with rather less resolution than is allowed by more invasive interventions in animals, such as local cooling, pharmacological manipulation or even genetic intervention in a specific brain region). Moreover, it has now become possible for the first time to combine TMS online with fMRI in human studies (Vuilleumier & Driver 2007), to study how manipulating activity in one specific brain region may influence others and to assess how this impacts causally on cognitive performance.

This brief survey shows that the past few decades have led to many remarkable advances in studies of brain function and of human cognition. But this Discussion meeting at the Royal Society, on Mental Processes in the Human Brain (held 16–17 October 2006), was not intended to provide a historical overview of how the field got here. Instead, we charged the speakers and contributors with surveying what is currently known, and what new challenges and opportunities arise for the foreseeable future. We were inspired by several prior Royal Society Discussion meetings on related topics (including Broadbent & Weiskrantz 1982; Roberts *et al.* 1996; Parker *et al.* 2002, among others). But, we deliberately set out to organize this particular meeting along somewhat different lines. The Broadbent & Weiskrantz (1982) meeting had focused on cognitive neuropsychology in patient studies, whereas here we deliberately interleave studies of normality with pathology. Roberts *et al.* (1996) focused primarily on the frontal lobe in particular, whereas we had no such restriction. Parker *et al.* (2002) focused primarily (but not exclusively) on physiological studies of cognitive function in animals, with some emphasis on sensory function. We focused instead on so-called higher-level cognitive functions

(e.g. memory, language, awareness, attention, executive function) in humans.

All these topics provide unequivocally 'cognitive' domains that feature prominently in human mental life, and that in some cases (e.g. for language) may have no direct animal homologue. Since no prior Royal Society discussion meeting had focused extensively on the advances, new possibilities and possible shortcomings of functional neuroimaging, we address these in some detail here. This seemed appropriate, as the advent of neuroimaging has provided arguably the biggest sea change in studies of human cognitive and brain function in recent years (albeit not always without its critics; see Coltheart 2006). We were not able to cover all of the recent developments in the field. For instance, there is relatively little here on the growth of so-called social neuroscience, nor on developmental aspects. Such aspects are covered elsewhere (e.g. Frith & Frith 2003; Emery *et al.* in press).

Hagoort & van Berkum (2007), Marslen-Wilson & Tyler (2007) and Patterson (2007) provide insights here into how the new methods and theories have influenced studies of human language function, including mental representation in the brain of semantics, syntax, morphemes and even of pragmatic contextual constraints during communication. D'Esposito (2007), Dolan (2007) and Schacter & Addis (2007) survey recent developments for different aspects of memory and learning. Burgess *et al.* (2007), Cohen *et al.* (2007), Robbins (2007) and Stuss & Alexander (2007) present advances in the study of so-called 'executive functions' (or top-down cognitive control), relating not only to frontal cortex, but also to the many systems that specific frontal regions interconnect with, and to pharmacological modulation of such loops (Robbins 2007; see also Dolan 2007). Kouider & Dehaene (2007), Rees (2007) and Vuilleumier & Driver (2007) report on recent studies of perceptual awareness and attention in the human brain. They highlight both theoretical (Kouider & Dehaene 2007) and methodological advances (Rees 2007; Vuilleumier & Driver 2007), in addition to several key findings.

Although the presentations from all these contributors were organized into four separate sessions at the meeting (on language, memory, awareness and attention and executive function), there is often much striking overlap between the subtopics. For instance, frontal cortex features not only in the executive functions topic, but also in the language contribution by Hagoort & van Berkum (2007); in Dolan's (2007) account of how learning and affect impact upon conditioned responses and decision making; and in the three contributions on awareness and attention (Kouider & Dehaene 2007; Rees 2007; Vuilleumier & Driver 2007). Equally, D'Esposito's (2007) contribution is arguably concerned as much with executive function as it is with short-term or working memory and so on. All of the contributions emphasize the need to go beyond just the particular contribution of each distinct brain area, to understand further how the various regions may interact causally in network terms, a topic that receives particular attention from Vuilleumier & Driver (2007).

There was much lively discussion at the meeting, which was the best attended ever in the history of Royal Society discussion meetings to date (with the audience spilling out into four overflow rooms!). We think that this exceptional attendance is a testament to the excitement and rapid rate of progress in this field, and to the intrinsic interest of our mental lives and their neural basis. All of the extended discussions that took place at the meeting have fed back into this volume.

There has been no better time to study the neural basis of human cognitive function. We hope that the present volume captures this, by illustrating the recent advances, excitement and future potential in this field.

We thank all participants at the discussion meeting; the speakers and contributors; Uta Frith FRS for chairing the language session; Jay McClelland for provocative comments; Rosalyn Lawrence from the UCL Institute of Cognitive Neuroscience, and Laura Howlett and many Royal Society staff for administrative help; James Joseph at the Phil. Trans. B editorial office; and our many colleagues at the UCL Institute of Cognitive Neuroscience and neighbouring centres in Queen Square, all of whom share our passion for studying mental processes in the human brain. We also thank participants at the separate Festschrift for Tim Shallice held at UCL on 18 October, subsequent to the Royal Society Discussion meeting. A video recording of the discussion meeting is available at: <http://www.royalsoc.ac.uk/page.asp?id=1110>

Jon Driver*
Patrick Haggard
Tim Shallice

May 2007

UCL Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1N 3AR, UK

*Author for correspondence (j.driver@ucl.ac.uk)

REFERENCES

- Broadbent, D. E., Weiskrantz, L. (eds). 1982 The neuropsychology of cognitive function. Proceedings of a Royal Society Discussion Meeting Held on 18 and 19 November 1981. *Phil. Trans. R. Soc. B* **298**(1089).
- Burgess, P. W., Gilbert, S. J. & Dumontheil, I. 2007 Function and localization within rostral prefrontal cortex (area 10). *Phil. Trans. R. Soc. B* **362**, 887–899. (doi:10.1098/rstb.2007.2095)
- Cohen, J. D., McClure, S. M. & Yu, A. J. 2007 Should I stay or should I go? How the human brain manages the tradeoff between exploitation and exploration. *Phil. Trans. R. Soc. B* **362**, 933–942. (doi:10.1098/rstb.2007.2098)
- Coltheart, M. 2006 What has functional neuroimaging told us about the mind (so far)? *Cortex* **42**, 323–331.
- Dayan, P. & Abbot, L. F. 2005 *Theoretical neuroscience*. Cambridge, MA: MIT Press.
- D'Esposito, M. 2007 From cognitive to neural models of working memory. *Phil. Trans. R. Soc. B* **362**, 761–772. (doi:10.1098/rstb.2007.2086)
- Dolan, R. J. 2007 The human amygdala and orbital prefrontal cortex in behavioural regulation. *Phil. Trans. R. Soc. B* **362**, 787–799. (doi:10.1098/rstb.2007.2088)
- Emery, N., Clayton, N. & Frith, C. 2007 Introduction. Social intelligence: from brain to culture. *Phil. Trans. R. Soc. B* **362**, 485–488. (doi:10.1098/rstb.2006.2022)
- Frith, U. & Frith, C. D. 2003 Development and neurophysiology of mentalizing. *Phil. Trans. R. Soc. B* **358**, 459–473. (doi:10.1098/rstb.2002.1218)

- Fuster, J. M., Bauer, R. H. & Jervey, J. P. 1985 Functional interactions between inferotemporal and prefrontal cortex in a cognitive task. *Brain Res.* **330**, 299–307. (doi:10.1016/0006-8993(85)90689-4)
- Hagoort, P. & van Berkum, J. 2007 Beyond the sentence given. *Phil. Trans. R. Soc. B* **362**, 801–811. (doi:10.1098/rstb.2007.2089)
- Hariri, A. R., Drabant, E. M. & Weinberger, D. R. 2006 Imaging genetics. *Biol. Psychiatry* **59**, 888–897. (doi:10.1016/j.biopsych.2005.11.005)
- Hinton, G. & Shallice, T. 1991 Lesioning an attractor network: investigations of acquired dyslexia. *Psychol. Rev.* **98**, 74–95. (doi:10.1037/0033-295X.98.1.74)
- Hubel, D. H. & Wiesel, T. N. 1959 Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* **148**, 574–591.
- Kandel, E. R. 2004 The molecular biology of memory storage: a dialog between genes and synapses. *Biomed. Life Sci.* **24**, 475–522.
- Kouider, S. & Dehaene, S. 2007 Levels of processing during non-conscious perception: a critical review of visual masking. *Phil. Trans. R. Soc. B* **362**, 857–875. (doi:10.1098/rstb.2007.2093)
- Logothetis, N. K. & Schall, J. D. 1989 Neuronal correlates of subjective visual perception. *Science* **245**, 761–763. (doi:10.1126/science.2772635)
- Marslen-Wilson, W. D. & Tyler, L. K. 2007 Morphology, language and the brain: the decompositional substrate for language comprehension. *Phil. Trans. R. Soc. B* **362**, 823–836. (doi:10.1098/rstb.2007.2091)
- McCarthy, R. A. & Warrington, E. K. 1988 *Cognitive neuropsychology*. London, UK: Academic Press.
- McClelland, J. L. & Rumelhart, D. E. 1985 Distributed memory and the representation of general and specific information. *J. Exp. Psychol. Gen.* **114**, 159–188. (doi:10.1037/0096-3445.114.2.159)
- Moran, J. & Desimone, R. 1985 Selective attention gates visual processing in the extrastriate cortex. *Science* **229**, 782–784. (doi:10.1126/science.4023713)
- Newsome, W. T. & Britten, K. H. 1989 Neural correlates of a perceptual decision. *Nature* **341**, 52–54. (doi:10.1038/341052a0)
- O'Keefe, J. & Dostrovsky, J. 1971 The hippocampus as a spatial map: preliminary evidence from unit activity in the freely moving rat. *Brain Res.* **34**, 171–175. (doi:10.1016/0006-8993(71)90358-1)
- Parker, A., Derrington, A. & Blakemore, C. 2002 The physiology of cognitive processes. *Phil. Trans. R. Soc. B* **357**, 959–961. (doi:10.1098/rstb.2002.1115)
- Patterson, K. 2007 The reign of typicality in semantic memory. *Phil. Trans. R. Soc. B* **362**, 813–821. (doi:10.1098/rstb.2007.2090)
- Posner, M. I. & Raichle, M. E. 1994 *Images of mind*. San Francisco, CA: Freeman and company.
- Rees, G. 2007 Neural correlates of the contents of visual awareness in humans. *Phil. Trans. R. Soc. B* **362**, 877–886. (doi:10.1098/rstb.2007.2094)
- Robbins, T. W. 2007 Shifting and stopping: fronto-striatal substrates, neurochemical modulation and clinical implications. *Phil. Trans. R. Soc. B* **362**, 917–932. (doi:10.1098/rstb.2007.2097)
- Roberts, A. C., Robbins, T. W. R. & Weiskrantz, L. 1996 Executive and cognitive functions of the prefrontal cortex. *Phil. Trans. R. Soc. B* **351**, 1389–1395. (doi:10.1098/rstb.1996.0122)
- Schacter, D. L. & Addis, D. R. 2007 The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Phil. Trans. R. Soc. B* **362**, 773–786. (doi:10.1098/rstb.2007.2087)
- Scoville, W. B. & Milner, B. 1957 Loss of recent memory after bilateral hippocampal lesions. *J. Neurol. Neurosurg. Psychiatry* **20**, 11–21.
- Shallice, T. 1988 *From neuropsychology to mental structure*. Cambridge, UK: Cambridge University Press.
- Stuss, D. T. & Alexander, M. P. 2007 Is there a dysexecutive syndrome? *Phil. Trans. R. Soc. B* **362**, 901–915. (doi:10.1098/rstb.2007.2096)
- Valdes-Sosa, P. A., Koptter, R. & Friston, K. J. 2005 Introduction: multimodal neuroimaging of brain connectivity. *Phil. Trans. R. Soc. B* **360**, 865–867. (doi:10.1098/rstb.2005.1655)
- Vuilleumier, P. & Driver, J. 2007 Modulation of visual processing by attention and emotion: windows on causal interactions between human brain regions. *Phil. Trans. R. Soc. B* **362**, 837–855. (doi:10.1098/rstb.2007.2092)
- Walsh, V. & Pascual-Leone, A. 2003 *Transcranial magnetic stimulation*. Cambridge, MA: MIT Press.