FROM MONKEY BRAIN TO HUMAN BRAIN

A Fyssen Foundation Symposium

EDITED BY

Stanislas Dehaene, Jean-René Duhamel, Marc D. Hauser, and Giacomo Rizzolatti

From Monkey Brain to Human Brain

Fyssen Foundation Series, published by Oxford University Press

Social Relationships and Cognitive Development
edited by Robert Hinde, Anne-Nelly Perret-Clermont, and Joan Stevenson-Hinde

Thought Without Language
edited by L. Weiskrantz

The Use of Tools by Human and Non-human Primates
edited by Arlette Berthelet and Jean Chavaillon

Origins of the Human Brain
edited by Jean-Pierre Changeux and Jean Chavaillon

Causal Cognition: A Multidisciplinary Debate
edited by Dan Sperber, David Premack, and Ann James Premack

From Monkey Brain to Human Brain
edited by Stanislas Dehaene, Jean-René Duhamel, Marc D. Hauser, and Giacomo Rizzolatti

From Monkey Brain to Human Brain

A Fyssen Foundation Symposium

edited by Stanislas Dehaene, Jean-René Duhamel, Marc D. Hauser, and Giacomo Rizzolatti

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

© 2005 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

MIT Press books may be purchased at special quantity discounts for business or sales promotional use. For information, please email special_sales@mitpress.mit.edu or write to Special Sales Department, The MIT Press, 5 Cambridge Center, Cambridge, MA 02142.

This book was set in Stone sans and Stone serif by SNP Best-set Typesetter Ltd., Hong Kong and was printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

From monkey brain to human brain : a Fyssen Foundation symposium / editors, Stanislas Dehaene . . . [et al.].
p. cm.—(Fyssen Foundation symposium)
"Proceedings of the ninth Fyssen Symposium entitled 'From monkey brain to human brain' which was held at the Pavillon Henry IV in St-Germain-en-Laye from 20 to 23 June 2003"—Pref.
Includes bibliographical references and index.
ISBN 0-262-04223-1 (alk. paper)
1. Brain—Evolution—Congresses. 2. Comparative neurobiology—Congresses. 3. Primates—Physiology—Congresses. I. Dehaene, Stanislas. II. Fyssen Symposium (9th : 2003 : St. Germain-en-Laye, France). III. Series.

QP376.F715 2005
612.8'2—dc22

2004051170

10 9 8 7 6 5 4 3 2 1

To the memory of Mrs. A. H. Fyssen

Contents

Series Foreword

The aim of the Fyssen Foundation is to "encourage all forms of scientific inquiry into cognitive mechanisms, including thought and reasoning, that underly animal and human behavior; their biological and cultural bases, and phylogenetic and ontogenetic development."

The Foundation aims to support research that will lead to a more rigorous and precise approach to this fundamental domain, calling upon such disciplines as ethology, paleontology, archaeology, anthropology, psychology, logic, and the neurosciences.

The Foundation is named after its founder, Mr. A. H. Fyssen, a French businessman who has long been interested in the scientific understanding of cognitive problems, and who is responsible for its endowment. Its headquarters are at 194, rue de Rivoli, 75001 Paris. The Foundation was recognized as a charitable institution by the French government by decree on March 20, 1979.

The Fyssen Foundation has developed a program to support research in the abovementioned area. This program includes in particular:

- Post doctoral study grants, for the training and support of scientists working in fields which coincide with the aims of the Foundation;
- Research grants, to support scientists in the field or in laboratories, who are working along lines of research corresponding to the objectives of the Foundation;
- The organization of its own symposia and publications on topics considered as important in fulfilling the aims of the Foundation;
- The regular publication of the *Fyssen Annals*, which include original articles in the fields supported by the Foundation;
- The Fyssen International Scientific Prize, given for a decisive contribution to the progress of knowledge in the fields of research supported by the Foundation. The prize is given each year to a scientist of international reputation.

The Foundation is administered by a Board of Directors consisting of scientists, lawyers, and financiers, and by an International Scientific Committee in charge of the scientific policy of the Foundation, of the launching of its programmes, of the scientific evaluation of the applications, of the assessment of the work supported by the Foundation, and of the election of the recipient of the Fyssen International Scientific Prize.

Preface

Looking at the extraordinary overlap between human and chimpanzee genomes, one might expect equally extraordinary overlap in their thoughts, sensations, perceptions, and emotions. Yet there are both considerable similarities and differences between human and nonhuman primate brains and minds. It is not yet clear how to capture these interesting patterns of convergence and divergence, especially in terms of identifying the relevant neural structures and the selective pressures that helped shape each species' psychological signature. In what ways are our minds and brains similar to those of other primates? And where do the critical differences lie? New advances in cognitive psychology, comparative biology, and neuroscience have created an opportunity to take a fresh look at these complex and fascinating problems. It therefore seemed timely to assemble a panel of distinguished researchers in those fields, with the goal of presenting a state-of-the-art comparative perspective on primate cortical organization and function. Such was the aim of the ninth Fyssen Symposium, entitled "From Monkey Brain to Human Brain," which was held at the Pavillon Henry IV in St-Germain-en-Laye from 20 to 23 June 2003, and of which the present volume constitutes the report.

The first part of the meeting was dedicated to an examination of the potentials and limits of the modern techniques for studying primate brains. Discussion focused largely on neuroimaging methods, but other approaches were also examined, including those provided by recent advances in genetics and computer-based reconstructions. These methods can now be applied identically across different species of primates, including humans, and may thus serve as a starting point for the definition of cross-species homologies.

Neuroimaging, in particular, offers a diversity of means to compare human and nonhuman primate brains. At the anatomical level, macroscopic images of brain volume or surface can be warped onto each other, thus allowing for precise measurements of the amount of regional distortion. A comprehensive approach to this problem is presented in David Van Essen's chapter. At the functional level, the response profile of different brain areas to the same stimuli can be measured and compared across species

using functional magnetic resonance imaging (fMRI). Ultimately, fMRI should provide a strong connection to single-cell recordings in the behaving monkey, although there are important caveats that are discussed by Zoe Kourtzi and Nikos Logothetis. Anatomical and functional MRI results can be integrated within computerized atlases that, increasingly, also incorporate detailed postmortem data. For instance, Karl Zilles describes how the regional and laminar distribution of various types of neurotransmitters and receptors provides a rich source of comparative data across species.

Karl Zilles and Jean-Jacques Hublin also examine the extent to which measures of brain size, shape, gyrification, and vascularization shed light on the evolution of the human brain and the major factors that shaped its organization. As described by Jean-Pierre Changeux, advances in genomics have merely strengthened, rather than resolved, the paradox of the nonlinearity of primate evolution, which is the fact that major changes in brain size and functional complexity resulted from small changes in the genome. This paradox raises many unsolved questions, which are explored in the first chapters:

- Was human brain evolution driven mostly by a change in cortical surface?
- What is the meaning of the ten- to twenty-fold variation in relative size between different cortical areas of the human and macaque brains?
- Does a single deformation field suffice to account for cross-species differences in cortical maps? Or is there also evidence that new cortical areas have appeared? Might even more radical forms of reshuffling of the cortical layout have occurred?
- Have basic units such as cortical layers and columns been preserved in evolution, or did their functionality change?
- Were there also major changes in microcircuitry, for instance in axonal arborization and long-distance connectivity? What is their significance?

A second section of the meeting examined particular domains of human competence and their possible precursors in primates. We intentionally focused first and foremost on cognitive functions that appear to be particularly developed in our species, to such an extent that they initially appear unique and devoid of precursors: arithmetic, reading, theory of mind, cooperation, and altruism. Remarkably, in all of these domains, some plausible analogs or even homologs have begun to be identified in nonhuman primates. In arithmetic, for instance, a coherent story is emerging that relates the human capacity for symbolic calculation to a more primitive form of numerosity processing, available to primates as well as to many other species including rats and pigeons. As reviewed by Elizabeth Brannon, behavioral studies in primates have revealed a competence for discriminating and comparing the numerosity of sets of dots. This competence for nonsymbolic arithmetic is also present in preverbal human infants, and may thus be part of our evolutionary heritage at birth. A very

exciting development is that its neurobiological bases in monkeys have begun to be identified. Based on electrophysiological recordings, Andreas Nieder and Earl Miller identify a population of single neurons tuned to approximate numerosity. They demonstrate that the properties of these neurons explain the distance effect and Weber's law that characterize monkey and human number-processing behavior. Using fMRI and anatomical warping methods, Stanislas Dehaene proposes that the parietal areas that have been identified as the substrates of numerosity processing in humans and macaques are plausible homologs. Only humans, however, have the ability to access this numerosity representation through written and spoken symbols, which allows them to develop a full-fledged system of exact arithmetic. More generally, Dehaene proposes that human cultural inventions such as reading and arithmetic rely on a "recycling" of existing cerebral areas (a minimal conversion to a novel use) rather than the *de novo* creation of cultural cortical circuits through an all-purpose learning system.

David Perrett and his colleagues tackle the issue of whether only humans have a capacity to understand the actions, intentions, and minds of others. Behavioral experiments, mostly from chimpanzees, converges with recent single-cell recordings in macaques to suggest that primates possess a remarkable degree of competence for inferring the intentions of other congeners from their actions, with neural substrates in the superior temporal sulcus. In related work, Giacomo Rizzolatti and Giovanni Buccino examine the power and limits of the parieto-frontal "mirror neuron" systems, which provides a joint representation of the actions of oneself and of others. They demonstrate that a homologous parietofrontal mirror system is present in humans, and speculate that this system, common to production and comprehension of actions, could have played a crucial role in the emergence of language. Giuseppe Luppino describes in great detail the anatomy of parietal and frontal connections on which the action system is founded. He presents new anatomical findings on visual connections of the posterior parietal cortex and discuss the homologies between monkey and human parietal lobe organization. Atsushi Iriki further describes how part of these circuits are modified when macaque monkeys learn to use a tool in order to reach for objects. This innovative behavioral paradigm, which lends itself to basic neurophysiological, neuroimaging, and gene expression experiments, provides a new standpoint from which to speculate about the evolution of human toolmaking ability.

The issue of action understanding is taken one step further by Jeffrey Stevens and Marc Hauser, who ask how primates evolved the ability to cooperate with one another. They outline a series of constraints that any biological species must meet in order to develop cooperation, and specifically reciprocal altruism. Furthermore, they describe a new behavioral task in which multitrial reciprocation develops or fails to develop between two tamarin monkeys as a function of whether the initial acts of cooperation are intentional or not, thus revealing some of the behavioral mechanisms of

cooperation in nonhuman primates. These studies open the door to future neurobiological assays of cooperation in human and nonhuman primates.

Cognitive control constitutes an essential element of cooperation as well as any other form of complex behavior extending over a period of time. Michael Petrides considers how the primate prefrontal cortex contributes to elaborate tasks that require cognitive control over the contents of working memory. Céline Amiez, Jean-Paul Joseph, and Emmanuel Procyk extend this analysis to the anterior cingulate cortex, which in humans appears as an essential element of many cognitive control tasks such as the Stroop test. They review anatomical, single-cell, and lesion studies that indicate that rewards and errors may be processed similarly by the anterior cingulate in both human and nonhuman primates.

Beyond these high-level cognitive functions, it is also essential, and perhaps simpler, to examine to what extent basic functions such as visual recognition and visual attention are also significantly related in humans and in other primates. Elinor McKone and Nancy Kanwisher review human neuroimaging evidence in favor of the hypothesis that a subpart of the ventral visual system houses mechanisms evolved for face recognition, as opposed to more general forms of expertise for within-category variation. This face recognition system may also be present in other primates. More generally, as discussed by David Van Essen, the primate inferotemporal cortex may constitute a plausible homolog of the human visual fusiform region for visual recognition, although the amount of cortical distortion and reorganization in this area remains to be fully understood. The inner workings of the primate inferotemporal cortex are analyzed in detail by Manabu Tanifuji, Kazushige Tsunoda, and Yukako Yamane. They describe convergent single-cell and optical imaging studies that begin to reveal how visual objects composed of multiple parts are coded by neural populations. Stanislas Dehaene argues that those populations of neurons may constitute an evolutionary precursor of the human reading system, since they require only minimal modification to be reconverted into an invariant visual letter and word recognition device.

Finally, Claire Wardak, Suliann Ben Hamed, and Jean-René Duhamel ask whether the macaque parietal lobe houses a visual attentional system comparable to the one observed in human neuropsychological and neuroimaging experiments. Parietal lesions in macaques lead to extinction and lesion orienting deficits comparable to those observed in human neglect patients, thus lending credibility to the hypothesis of a strong cross-species homology at this level, too.

The analysis of precursors of human abilities raises many further questions that cut across domains of cognitive competence. When does evidence of behavioral homology imply neurobiological or computational homology? To what extent do uniquely human cognitive abilities such as mathematics and reading rely on evolutionary ancient mechanisms? How can work on the evolutionary function of particular cog-

nitive abilities integrate with work at the level of mechanism? None of the chapters provide any definitive answers to these difficult issues. Yet, by giving some concrete examples of possible homologies between human and nonhuman primates, this book as a whole sets the stage for any further attempts to characterize human nature.

In closing this preface, we would like to express our profound gratitude to the Fyssen Foundation for making this symposium possible. The topic of the symposium coincided nicely with the aim of the Fyssen Foundation, which is "to encourage all forms of scientific inquiry into cognitive mechanisms, including thought and reasoning, underlying animal and human behaviour, and their ontogenetic and phylogenetic developments." The symposium was organized by two members of the Scientific Committee of the Foundation, Stanislas Dehaene and Giacomo Rizzolatti, with the help and advice of a previous member, Marc Hauser, and an external expert, Jean-René Duhamel (who was also one of the first to receive a Fyssen fellowship in 1985).

Unfortunately, a shadow was cast on our scientific discussions by the sad news of the death of Madame Fyssen, only a few days before the meeting. Since 1982, Madame Fyssen had been the president of the Foundation, which she had created together with her husband in 1979. In spite of her age, she attended every meeting of the Scientific Council, and she had been looking forward to this symposium. Her personality and presence were driving forces that helped maintain the Foundation's focus throughout the years. Thanks to her generosity, the Foundation continues to play an important role in the development of brain and cognitive sciences, particularly in France, by awarding fellowships and grants to young researchers as well as a renowned international prize.

We are extremely grateful to the staff of the foundation, Nadia Ferchal, Fanny Bande and Julie Rubin for their efficacious support of both the organization of the meeting and the publication of this book.

I Human Brain Evolution: New Methods and Results

1 Surface-Based Comparisons of Macaque and Human Cortical Organization

David C. Van Essen

In his pioneering architectonic studies of primate cerebral cortex, Brodmann (1909) described a rich mosaic of anatomically distinct cortical areas in both humans and monkeys. He identified 28 neocortical areas in the Old World monkey and 46 in the human, and he used corresponding terminology for most of these areas on the supposition that the similarities in architecture reflected evolutionary homologies. Subsequent studies leave little doubt that the primary sensory and motor areas and their immediate neighbors are indeed homologous in monkeys and humans. On the other hand, the evolutionary relationships are much less clear for most of the remaining expanse of neocortex, mainly because Brodmann's partitioning schemes for both species have been subject to extensive revision over the ensuing century. These revisions are based on many different lines of experimental evidence and are very much a work in progress; consensus has yet to be reached regarding the basic arrangement of cortical areas over most of neocortex in either species. Depending on the criteria used for identifying areas (itself a contentious issue), the total number of cortical areas may approach or exceed 100 areas in the macaque and an even larger number in humans, i.e., double or triple the number enumerated by Brodmann (Van Essen, 2004).

Because human cortex has 10 times the surface area of macaque cortex and plays a key role in many distinctive aspects of human cognition, there presumably are major species differences in cortical functional organization. A priori, these differences might reflect any combination of four basic possibilities:

1. Bigger areas Some areas may have increased in size in humans compared to macaques, thereby providing greater processing power for whatever functions they mediate.
2. Functional divergence Functional specialization of cortical areas may have undergone evolutionary divergence, such that the tasks mediated by homologous cortical areas may be very different in humans compared to macaques.
3. Areas gained or lost Completely new areas may have emerged along the human evolutionary trajectory, analogous to the gene duplication that has often occurred

during evolution of the genome. Alternatively, areas present in a common ancestor may have disappeared in one species but not the other.

4. Rearrangements Topological rearrangement of cortical areas (analogous to "jumping genes" in chromosomal DNA) may have occurred along one evolutionary trajectory but not the other.

To distinguish among these possibilities requires accurate maps of cortical organization in each species plus objective methods for making comparisons between maps. A fundamental challenge in mapping the cortex arises from cortical convolutions—both their existence and the dramatic species differences in the pattern of convolutions. Despite its convolutions, the cortex is a continuous sheet of tissue, topologically equivalent to a disc, and it can be represented by explicit surface reconstructions that capture the intricacies of cortical shape. Surface reconstructions facilitate visualization of many aspects of cortical organization that are difficult to decipher when viewing a series of slices through the brain. Moreover, the differences in cortical shape can be eliminated by mapping each cortical surface to a standard configuration, such as a sphere. One sphere can then be registered to another, constrained by landmarks that reflect known or suspected homologies. Consequently, surface-based registration provides a general and powerful strategy for analyzing species differences in cortical organization.

This chapter illustrates how surface-based visualization and interspecies registration can help clarify a number of specific issues and controversies regarding the functional organization of human and macaque cerebral cortex. The analysis is focused on two sets of areas situated at opposite ends of the hemisphere: orbital and medial prefrontal cortex (OMPFC) and visuotopically organized portions of occipital visual cortex. These choices are based on the availability of detailed maps of cortical organization in both regions for both species obtained using modern experimental approaches.

To set the stage for this analysis, figure 1.1 shows surface reconstructions of macaque and human right cerebral hemispheres, generated from high-resolution structural MRI data using the SureFit segmentation method and visualized using Caret software (Van Essen et al., 2001, 2004). The surfaces are displayed in five standard configurations; the shading on each map represents cortical depth (deeper is darker), which provides a convenient measure of the original cortical shape. The fiducial surfaces (panels A, F) represent the shape of the cortex, including all of the convolutions. The inflated maps (panels B, G) retain the approximate shape of the brain but smooths out all but the deepest folds. The spherical maps (panels C, H) provide a geometrically precise representation that is the substrate for registration between species. It also provides the basis for surface-based coordinates that concisely and objectively specify locations on the cortical surface, as indicated by the latitude (black) and longitude (gray)

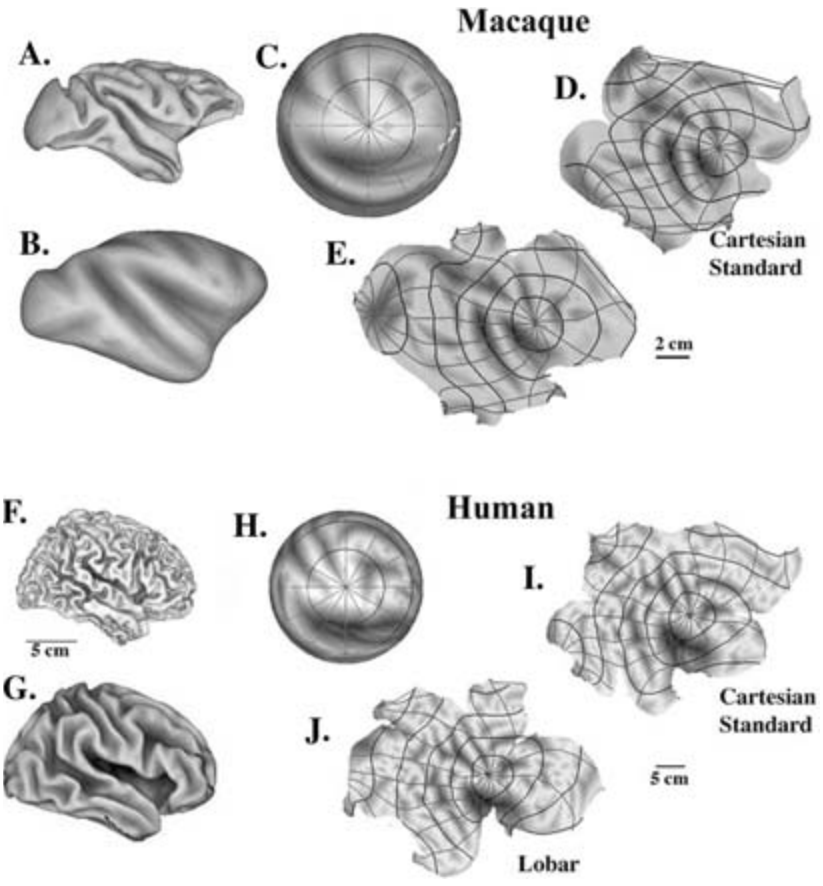


Figure 1.1

Surface-based atlases of human and macaque cortex. (*A–E*) Right cerebral hemisphere of the macaque F99UA1 atlas (Van Essen et al., 2004). (*F–J*) Right cerebral hemisphere of the human Colin atlas (Van Essen, 2002). Both atlas surfaces are shown in five configurations: fiducial, inflated, spherical, Cartesian standard flat, and lobar flat. The spherical and flat maps include latitude and longitude isocontours used for defining spherical coordinates.

isocontours on each spherical map in figure 1.1. The flat maps allow the entire cortical sheet to be seen in a single view without severe distortions (akin to flat maps of the earth's surface). Panels D and I show the commonly used Cartesian standard configuration; panels E and J show the "lobar" configuration that is better suited for the data analyzed here because it avoids cuts in occipital and frontal lobes. Each of the flat maps contains a different pattern of areal distortions relative to the fiducial surface. Various differences that are discussed below regarding the relative sizes of particular regions and areas are based on surface area measurements of the fiducial surface, not on the sometimes deceptive surface areas on the flat maps.

The macaque atlas map in figure 1.2A–D (see also plate 1) shows visuotopically organized areas in occipital cortex and posterior temporal parietal cortex, as identified in the Felleman and Van Essen (1991) partitioning scheme. The human atlas map (figure 1.2F–I) includes visuotopic areas from fMRI mapping studies (Hadjikhani et al., 1998; see Van Essen, 2004). In addition, panels E and J show alternate schemes for ventral occipitotemporal cortex in macaque and human. Both atlases include maps of architectonic areas in orbital and medial prefrontal cortex (OMPFC), identified using a combination of cytoarchitecture, myeloarchitecture, and immunocytochemistry (Carmichael & price, 1994; Ferry, Öngür, An, & Price, 2000; Öngür and Price, 2000; Öngür, Ferry, & Price, 2003). The atlas configurations include lateral and medial views of the fiducial surface (figure 1.2, A, B, F, G), inflated maps viewed from an anteroventral perspective (figure 1.2 C, H), and flat maps in the lobar configuration to avoid cuts where the areas have been mapped (figure 1.2 D, I).

All of the labeled regions shown in figure 1.2 differ from one another in significant respects, but not all of them are generally accepted as genuine cortical areas. In the terminology used here (see also Lewis and Van Essen, 2000), a cortical area refers to a well-defined region identifiable by one or more attributes that both unify the region and distinguish it from surrounding regions. A zone signifies a region in which one or more consistent regional differences have been reported, but may not warrant consideration as separate areas. A subdivision is a more neutral term, signifying a noncommittal label as to whether the region is an area or a zone.

The cortical areas shown in figure 1.2A–I were initially charted on individual hemispheres that had been analyzed using anatomical or functional methods. They were registered to the atlas maps using surface-based registration (Van Essen et al., 2001; Van Essen, Harwell, Hanlon, & Dickson, 2004), with geographic landmarks as constraints for the registration. Owing to the well-known individual variability in the location of areal boundaries relative to nearby geographic landmarks, there is inherently some uncertainty associated with the location of all areas on the atlas maps.

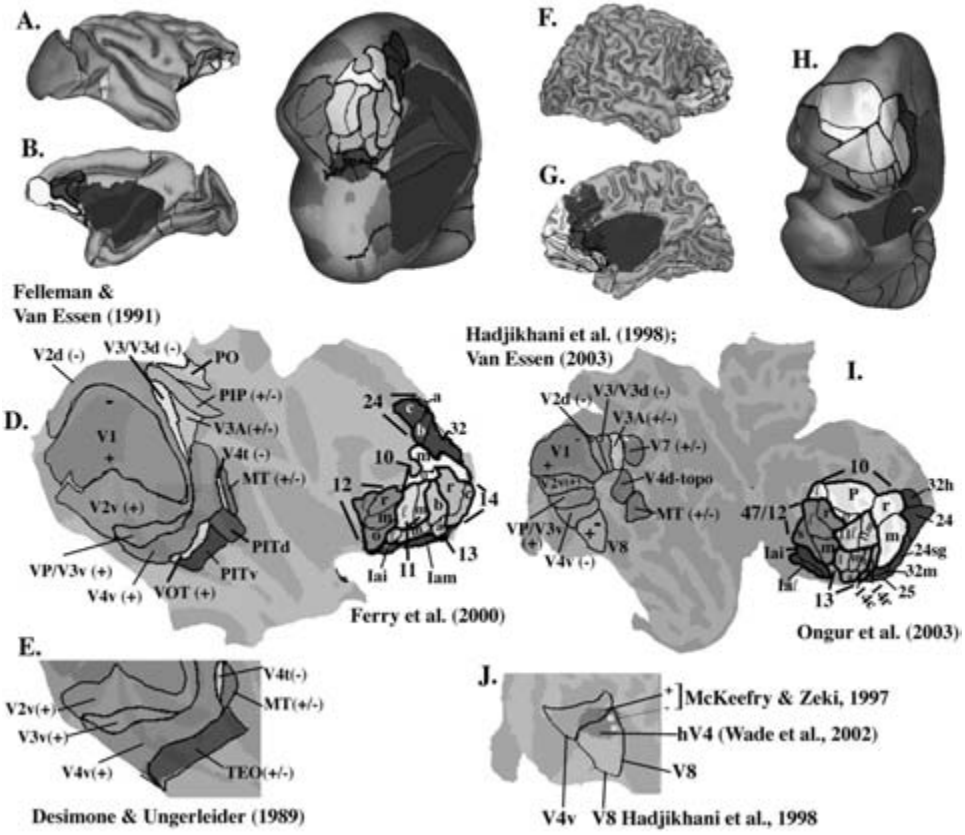


Figure 1.2

Visuotopic and orbitomedial prefrontal cortex (OMPFC) subdivisions of macaque and human cortex. (*A–D*) Fiducial (lateral and medial), inflated, and lobar flat map views of the macaque atlas with visuotopic areas (Felleman & Van Essen, 1991) and OMPFC areas (Ferry et al., 2000, case om 43). (*E*) Visual areas from Desimone and Ungerleider (1989) on a flat map of just ventral occipitotemporal cortex. (*F–I*) Fiducial, inflated, and lobar fat map views of the human atlas with visuotopic areas (Hadjikhani et al., 1998; see Van Essen, 2004) and OMPFC areas (Öngür et al., 2003; composite map generated as an average of three individual right hemispheres). (*J*) Human V4 as delineated by Wade et al. (2002, their figure 9b, case, A.W. right hemisphere) and by McKeefry and Zeki (1997; red, center of upper-field activation; green, center of lower-field activation). The visuotopic maps were registered using a 2-D registration algorithm applied to published images of flat maps. The OMPFC maps were registered by mapping the prefrontal surface reconstructions to a partial sphere, then registering this to the atlas sphere. In all cases, geographic (sulcal) landmarks were used to constrain the registration. Data sets used in generating this figure and figures 1.1–1.4 can be accessed via http://brainmap.wustl.edu:8081/sums/archivelist.do?archive_id=636599. See plate 1 for color version.

OMPFC Areas

In the macaque, Ferry et al. (2000) charted 20 orbitofrontal areas, as shown in figure 1.2A–D for one individual case mapped to the atlas map (see figure legend for details). As indicated by the coloring scheme and by the thicker borders around each area complex, most of these involve finer-grained subdivisions of Brodmann's numbering scheme, as modified by Walker (1940) and Petrides and Pandya (2002). In human cortex, Öngür et al. (2003) charted 24 orbitofrontal subdivisions, shown in figure 1.2B after mapping to the atlas. In general, there are many similarities in the layout of areas in the two species, but some major differences as well. The most lateral cluster (left on the flat maps), includes four subdivisions of macaque area 12 and of human area 47/12, with both sets colored red to reflect the presumed homologies. More medial and ventral are clusters that include two subdivisions of area 11 (green), four of area 13 (light blue), and two of area 14 (orange). These differ in relative size (e.g., 14r and 14c are much smaller on the human map). Anterior and more dorsal area is the area 10 complex (yellow), whose five subdivisions in humans (10p, 10o, 10r, 10m, and 10l) occupy 4.5 percent of neocortical surface areqa, which is three-fold greater than the 1.4 percent occupied by the two subdivisions (10m, 10o) in the macaque. Medially is a complex of areas that includes subdivisions of areas 24, 25, and 32. The topological (neighborhood) relationships between different areas are generally similar in macaque and human. There are a few minor differences, comparable to the differences in individual hemispheres mapped within the same species (Ferry et al., 2000); it remains to be determined whether this reflects genuine variability in map topology versus experimental uncertainties in charting areal boundaries.

Visuotopic Subdivisions

In the macaque, cortex that is predominantly or exclusively visual in function occupies more than half of the total cortical surface area. There is evidence for up to 40 visual subdivisions (areas plus zones; Lewis and Van Essen, 2000) but considerably fewer in various other partitioning schemes (see Van Essen, 2004). Orderly visuotopic maps occur in many visual areas, particularly in occipital cortex. Figure 1.2 shows 16 visuotopic subdivisions of the Felleman and Van Essen (1991) scheme, in which the visuotopic maps progress from extremely precise and fine-grained in area V1 to very coarse in the posterior inferotemporal complex (PITd and PITv). Area V1 contains a complete map of the contralateral visual hemifield and is bounded by a representation of the vertical meridian. Area V2 shares the vertical meridian representation with V1 and includes a split representation of the horizontal meridian along the anterior boundaries of its upper field (+) and lower field (–) representations. Of the remaining visuotopic subdivisions, some have complete representations (indicated by +/– on the

flat map) but other representations are incomplete (+ or – on the map). Whether the partial-field representations constitute distinct visual areas is controversial (see below). The human map includes 11 visuotopically organized subdivisions, including several partial-field representations. The coloring scheme indicates potential correspondences between macaque and human, but not all of these necessarily represent genuine homologies.

The three clearest homologies are for areas V1, V2, and MT. In both species, V1 is the largest single area, but as a fraction of total cortex it is several times larger in the macaque than human cortex (10 percent vs. 3 percent). V2 is the second-largest area in both species. MT (also known as V5) is a much smaller area, distinguished by a high incidence of direction selectivity in the macaque (Van Essen et al., 1981) and by motion-selective PET and fMRI activations in humans (Watson et al., 1993; Hadjikhani et al., 1998). The map of the human motion-specific focus is identified as MT+ because it likely includes some of the adjoining motion-responsive MST complex. In both species MT has a similar visuotopic organization (Van Essen, Maunsell & Bixby, 1981; Huk, Dougherty, & Heeger, 2002).

In both the human and the macaque, V2 is adjoined dorsally by a lower-field representation and ventrally by an upper-field representation, referred to here as V3d and V3v respectively, rather than the alternate nomenclature of V3 and VP. In the macaque, V3d and V3v are reported to differ in some aspects of architecture, function, and connectivity (Van Essen, Newsome, Maunsell, & Bixby, 1986), though the magnitude of these dorsoventral asymmetries is controversial (Lyon & Kaas, 2002). Recent fMRI studies support the hypothesis of functional asymmetries between V3d and V3v (Tsao et al., 2003; Denys et al., 2003), but more detailed analyses are needed to assess the magnitude, nature, and significance of such asymmetries. The issue of whether V3d and V3v are separate areas or subdivisions of a unified V3 is to a large extent semantic, and the debate could be regarded as a tempest in a teapot if it applied only to V3d and V3v. However, analogous issues arise in the analysis of V4 and adjoining regions (see below), making the conceptual distinction of greater import.

In the macaque, both V3d and V3v are generally narrower than V2 when charted anatomically and neurophysiologically, consistent with their coarser visuotopic organization and larger receptive field sizes (Van Essen et al., 1986; Gattass, Sousa, & Gross, 1988). In contrast, fMRI-based estimate suggest that V3d and V3v are comparable in width to V2, both in the macaque (Brewer, Press, Logothetis, & Wandell, 2002; Fize et al., 2003) and in humans (figure 1.2B; Hadjikhani, Liu, Dale, Cavanagh, & Tootell, 1998; Wade, Brewer, Rieger, & Wandell, 2002; Dougherty et al., 2003). However, the fMRI-based estimates of areal boundaries in both species may be significantly biased as a consequence of the limited spatial resolution of fMRI with current methodology, and such biases could have a significant impact on estimated areal dimensions and surface areas.

V3A in both macaque and human involves a complete upper and lower field representation, albeit coarser and irregular. In the macaque V3A is adjoined medially by areas PIP and PO (Colby, Gattass, Olson, & Gross, 1988). In humans V3A is adjoined dorso-anteriorly by area V7 (Press et al., 2001).

In the macaque, area V4 includes a dorsal lower-field representation and a ventral upper-field representation that have been mapped physiologically (Gattass et al., 1988; Boussaud et al., 1991) and by fMRI (Brewer et al., 2002; Fize et al., 2003). Fize et al. (2003) describes a visuotopic asymmetry, in which the horizontal meridian representation forms the anterior boundary of V4 ventrally but not dorsally. V4t is a narrow strip lying between dorsal V4 and MT (Gattass et al., 1988) that represents lower fields, but it has not been resolved using fMRI. VOT is a narrow upper-field representation that has been mapped neurophysiologically (Van Essen et al., 1990; see also Boussaud et al., 1991) and by callosal connectivity (Van Essen et al., 1982) and fMRI mapping (Brewer et al., 2002, their figure 14). It lies anterior to V4v and posterior to the posterior inferotemporal complex, which includes two subdivisions (PITd and PITv) that each have a crude representation of upper and lower fields (Van Essen et al., 1990). In contrast, Boussaud et al. (1991) described TEO as a subdivision that subsumes VOT plus part of the adjoining PIT complex (figure 1.2E).

The location and nature of human area V4 remains controversial, with conflicting and at times confusing views regarding facts, terminology, and interpretation. Several studies have mapped an upper-field representation and identified it as human V4v because it lies in a corresponding location just anterior to V3v (i.e., a "topolog" of macaque V4v) and has a similar visuotopic organization (Sereno et al., 1995; DeYoe et al., 1996: Hadjikhani et al., 1998). These studies did not find a corresponding map of lower fields that would qualify as V4d. Hadjikhani et al. (1998) charted a separate representation of upper and lower fields that they identified as V8, lying antero-lateral to V4v (centered on [–38°, –122°] latitude and longitude on the atlas map). The foveal representation of V8 was clearly separate from that for V4v in one case, but the eccentricity mapping was ambiguous in other cases. The general region dorsolateral to V4v and posterior to MT has been variously identified as KO, LO, V3B LOC/LOP, or V4d-topo (Van Oostende et al., 1997; Smith et al., 1998; Tootell and Hadjikhani, 2001; Tsao et al., 2003). The V4d topolog (V4d-topo) name reflects its location relative to V2, V3A, V4v and MT, but its visuotopic organization is crude and does not match that of macaque V4d, nor is it mirror-symmetric to human V4v (Tootell and Hadjikhani, 2001). Thus, for both species it is an open question whether V4v and V4d/V4d-topo should be regarded as distinct areas or asymmetric components of a single area.

An alternative scheme (figure 1.2J) posits that human area V4 is a color-specific area restricted to ventral occipito-temporal cortex (Lueck et al., 1989; McKeefry & Zeki, 1997). Based on the Talairach stereotaxic coordinates of PET activation centers, its

upper-field representation maps to [–40°, –135°] on the atlas map (red in figure 1.2J) and its lower-field representation maps to [–34°, –135°] (green in figure 1.2J) with both foci close to the boundary of V8/V4v of Hadjikhani et al. (1998). However, in contrast to the situation with MT, evidence for a human color-specific activation provides only weak support for a homology with macaque V4 because macaque V4 is not specialized for color processing in the same way that MT is specialized for motion processing (Girard, Lomber, & Bullier, 2002; Cowey et al., 2001; see Felleman and Van Essen, 1991). Wade et al. (2002) mapped a representation of upper and lower fields in the same general region (blue in figure 1.2J for one of their individual cases mapped to the atlas). They consider the lower-field representation to be part of a single area, hV4, whose upper field includes V4v but not the upper-field component of V8 in the Hadjikhani et al. (1998) scheme. This interpretation is in accord with the McKeefry and Zeki (1997) scheme, but the data appear to be consistent also with the Hadjikhani et al. (1998) scheme for V8, given the noisiness and mapping uncertainties in the published data. Altogether, there is a pressing need for accurate, higher-resolution visuotopic maps in order to address the ambiguities and apparent discrepancies across studies. In the meantime, though, valuable additional insights can be obtained by comparing the published maps more closely using surface-based registration.

Surface-Based Registration

A key to interspecies registration is to identify a set of landmarks that can be reliably identified in both atlas maps and are highly likely to reflect genuine evolutionary homologies. The landmarks indicated in figure 1.3A and B (see plate 2) include early visual areas (V1, V2, and MT), other primary sensory areas (A1, olfactory, and gustatory cortex and the border between areas 3 and 4), the hippocampus, the olfactory sulcus, and additional landmarks along the natural boundary of cortex on the medial wall of the hemisphere. The relative positions of these landmarks (figure 1.3A–B) imply that highly nonuniform scaling must occur in several regions in order to achieve registration between the two maps. For example, V1 and V2 are a much smaller fraction of human compared to macaque cortex; the gap between MT and A1 is much larger on the human than the macaque map, and the gap between the frontal eye fields (FEF) and the boundary between somatosensory and motor cortex (areas 3 and 4) is much smaller on the human than the macaque map. These interspecies differences in relative location of functionally based landmarks greatly exceed the spatial uncertainties associated with each of the landmarks on the atlas maps, even for landmarks such as the FEF that are difficult to delineate with great accuracy in humans.

The landmark borders in figure 1.3 were drawn on flat maps and projected to the spherical maps. Registration was then carried out using an algorithm that deforms one spherical map to another, bringing the macaque landmarks into register with the

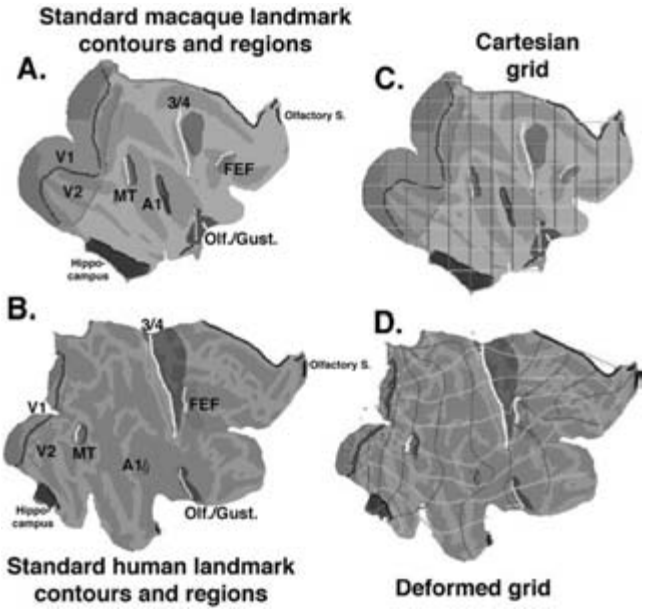


Figure 1.3
Landmarks used for registration between macaque and human right hemispheres. (*A*) Landmark areas and boundaries on the macaque atlas flat map (Cartesian standard). (*B*) Corresponding areas and boundaries on the human atlas flat map. (*C*) Cartesian grid on the macaque flat map. (*D*) Deformed macaque grid lines on the human map. (Reproduced with permission from Van Essen et al., 2004.) See plate 2 for color version.

human landmarks while minimizing shear and areal distortion in the intervening regions (Van Essen et al., 2004). A Cartesian grid on the macaque flat map (figure 1.3C) was projected to the macaque sphere, passively deformed to the human sphere, and projected to the human flat map (figure 1.3D). As expected from the relative locations of landmarks, the deformed grid is relatively compressed in occipital cortex and in posterior frontal cortex, whereas it is greatly expanded over much of parietal, temporal, and frontal cortex. For any given pair or triplet of landmarks, the registration algorithm results in relatively uniform expansion in the intervening region. If this results in good alignment between monkey and human areas that are known or suspected to be homologous in these intervening regions, then there is no need to invoke additional landmarks. If, on the other hand, the correspondence is poor, then additional or alternate landmarks can be explored.

Figure 1.4 (plate 3) shows the deformed macaque visuotopic and orbitofrontal areas, with the boundaries of the human areas overlaid. These are displayed on ventral, lateral, and medial views of the inflated surfaces (figure 1.4A–C), on a Cartesian stan-

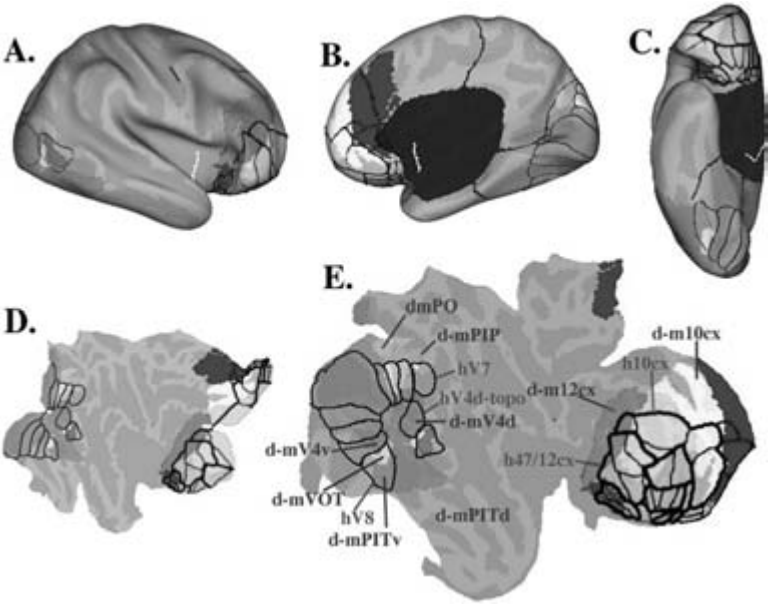


Figure 1.4
Deformed macaque areas (painted on surface) plus human areal boundaries (black contours) on the human atlas surface. (*A–C*) Ventral, lateral, and medial views of inflated configuration. (*D*, *E*) Cartesian standard and lobar flat map views. The prefix "d-m" signifies deformed macaque and is indicated by green labels for selected visual subdivisions (d-mVOT, etc.) and for OMPFC area complexes (d-m12cx, etc.). Selected human areas are identified by red labels and an "h" prefix (hV8, etc.). See plate 3 for color version.

dard flat map (figure 1.4D), and on a lobar cut flat map (figure 1.4E). In general, the deformed macaque OMPFC area complexes (d-m12cx, d-m10cx, etc.) occupy a substantially larger expanse on the human map (18 percent of total fiducial surface area) than do the corresponding human areas (about 11 percent of total surface area). Their expanded position occupies a broad swath of posterior and dorsal prefrontal cortex that includes portions of human areas 9, 45, 46, and 47. This mismatch between human and deformed macaque OMPFC areas strongly suggests a nonuniform expansion of prefrontal cortex in humans compared to macaques, in which dorsolateral and dorsomedial PFC expanded more than the OMPFC areas. This hypothesis can be made explicit by incorporating additional constraints, based on the boundaries of the OMPFC areas themselves. More generally, landmarks for homologies can be incorporated wherever a differential expansion of cortical regions is known or suspected. Of course, another option is to hypothesize that some of the proposed OMPFC areal homologies between human and macaque are not valid and to explore alternative candidate homologues suggested by other studies.

In visual cortex, deformed macaque V1, V2, and MT align well with their respective human counterparts, as, expected because these areas were used as landmarks. Deformed macaque V3d and V3v are narrower than their human counterparts, which may in part reflect an artifactual overestimate of the width of human V3d/v (see above). Deformed macaque V3A overlaps significantly with human V3A, consistent with their presumed homology. In contrast, human V7 does not overlap with deformed macaque PIP or PO, suggesting that human V7 lacks a known visuotopically organized homolog in the macaque and that macaque PIP and PO lack known visuotopically organized homologues in humans.

In ventral and lateral occipital cortex, deformed macaque V4v and V4d overlap extensively with human V4v and V4d-topo respectively, consistent with the homologies proposed by Tootell and Hadjikhani (2001). Deformed macaque VOT and PITv (but not PITd) lie mainly within human V8. In contrast, deformed macaque V4d (centered at [–24°, –134°]) longitude and latitude is very distant from human V8 and from the lower-field representation of human V4 (centered at [–33°, –134°]) proposed by McKeefry and Zeki (1997) and Wade et al. (2002). Hence, in order for this proposed homology to be valid, it would be necessary to invoke either (1) the emergence of a large cortical domain in human occipital cortex (lying between V2d, V3d, V3A, MT, and V4v) that has no homolog in the macaque, or (2) a major rearrangement in the topological relationships of homologous areas in the two species. While not impossible, neither of these possibilities is as plausible as the proposed homology between human V4d-topo and macaque V4d.

Comparing Macaque and Human Cerebellum

The cerebellum provides an interesting substrate for demonstrating the generality of surface-based interspecies comparisons, because cerebellar cortex is another sheetlike structure whose morphology and functional organization differs in many ways from cerebral cortex. Although cerebellar cortex is thinner and even more convoluted than cerebral cortex, it has recently been possible to generate accurate surface reconstructions of the full set of cerebellar lobules and lamellae and many of its fine-grained folia in both the human and macaque atlases (Van Essen, 2002). Figure 1.5 shows these cerebellar surface reconstructions in fiducial, spherical, and flat map configurations for macaque (figure 1.5A) and human (figure 1.5C). The cerebellar lobules are indicated by roman numerals alongside each flat map. As in figure 1.1, the shading on each map represents depth below the external hull. The elongation of the flat maps (particularly in the macaque) reflects the parallel folds along the cerebellar midline (center strip) and of the cerebellar hemispheres.

Figure 1.5B shows the results of deforming from macaque to human cerebellum using lobular boundaries as landmarks and spherical registration to constrain the

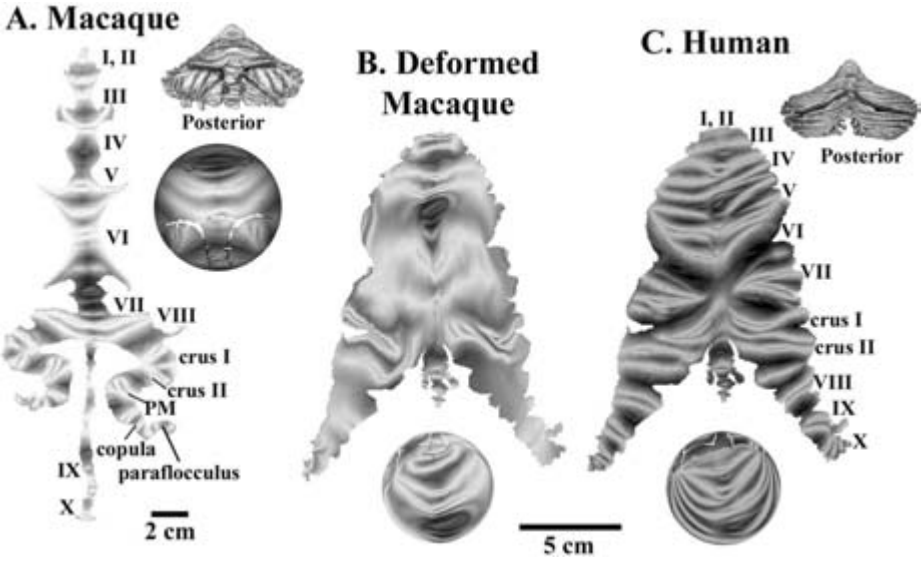


Figure 1.5
Macaque and human cerebellar maps. (*A*) Fiducial, spherical, and flat maps of macaque cerebellar cortex, with shading representing depth below the external hull of the cerebellum (lobules I–X and other geographic landmarks are shown on right). (*B*) Deformed macaque depth map after registration to the human spherical map (below) and flat map (above) using lobule boundaries to constrain the registration. (*C*) Fiducial, spherical, and flat maps of human cerebellar cortex. Data sets are accessible via http://brainmap.wustl.edu:8081/sums//archivelist.do?archive_id=632425.

deformation. Using these anatomical landmarks results in a pattern of differential expansion that is considerably greater than for cerebral cortex. Although the amount of experimental data on the cerebellar atlases is currently far less than for cerebral cortex, this approach will facilitate a wide variety of comparisons. For example, connectivity data obtained in the macaque (e.g., Kelly & Strick, 2003) can be mapped to the atlas, deformed to the human map and compared to fMRI data that have been mapped to the human cerebellar atlas. This should reveal whether functionally based subdivisions map to corresponding lobules in the macaque and human cerebellum.

Extending the Comparisons

The analyses and interspecies comparisons presented in this chapter can be extended in an open-ended way to all regions of cerebral and cerebellar cortex. They can be used to evaluated candidate homologies involving a wide variety of partitioning

schemes and many types of neuroimaging and other experimental data. Recent advances in brain-mapping software and databases allow this to be done in a flexible and efficient way. The current macaque and human surface-based atlases contain extensive data besides that illustrated in this chapter, including multiple partitioning schemes (14 for the macaque, 3 for human), connectivity and neurophysiology data in the macaque, and fMRI data (especially for human). The atlas data sets are accessible via the SumsDB database (http://brainmap.wustl.edu:8081/sums). Caret surface visualization software is freely available (http://brainmap.wustl.edu/caret) and runs on standard PC and Mac workstation platforms. This software also provides tools for mapping additional data to the atlas and entering data into the database. The specific data illustrated in this chapter can be downloaded from SumsDB (see URLs in figure 1.2 and figure 1.5 legends) and viewed in Caret.

In short, the stage is set for a fresh approach to studying human and macaque cortical organization that can capitalize on the explosion of experimental data being generated for both species. This provides an exciting opportunity to elucidate major commonalities and the nature of species differences that make us uniquely human.

Acknowledgments

Supported by a Human Brain Project/Neuroinformatics research grant funded jointly by the National Institute of Mental Health, National Science Foundation, National Cancer Institute, National Library of Medicine and the National Aeronautics and Space Administration (R01 MH60976-06), and The National Partnership for Advanced Computational Infrastructure (NPACI). I thank Dr. J. Price for comments on the manuscript, Susan Danker for assistance with manuscript preparation, and John Harwell, Donna Hanlon, James Dickson, and Heather Drury for development of Caret, SureFit, and SumsDB software.

References

Boussaud, D., Desimone, R., & Ungerleider, L. G. (1991). Visual topography of area TEO in the macaque. *Journal of Comparative Neurology*, *306*, 554–575.

Brewer, A. A., Press, W. A., Logothetis, N. K., & Wandell, B. A. (2002). Visual areas in macaque cortex measured using functional magnetic resonance imaging. *Journal of Neuroscience*, *22*, 10416–10426.

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Leipzig: J. A. Barth.

Carmichael, S. T., & Price, J. L. (1994). Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *Journal of Comparative Neurology*, *346*, 366–402.

Colby, C. L., Gattass, R., Olson, C. R., & Gross, C. G. (1988). Topographic organization of cortical afferents to extrastriate visual area PO in the macaque: a dual tracer study. *Journal of Comparative Neurology*, *269*, 392–413.

Cowey, A., Heywood, C. A., & Irving-Bell, L. (2001). The regional cortical basis of achromatopsia: a study on an achromatopsic patient. *European Journal of Neuroscience*, *14*, 1555–1566.

Denys, K., Vanduffel, W., Fize, D., Nelisen, K., Peuskens, H., Van Essen, D., & Orban, G. A. (2003). The processing of visual shape in the cerebral cortex of human and non-human primates: An fMRI study. In preparation.

Desimone, R., & Ungerleider, L. (1989). Neural mechanisms of visual processing in monkeys. In F. Boller & J. Graman (Eds.), *Handbook of Neuropsychology* (pp. 267–299). Amsterdam: Elsevier.

DeYoe, E. A., Carman, G., Bandetinni, P., Glickman, S., Wieser, J., Cox, R., Miller, D., & Neitz, J. (1996). Mapping striate and extrastriate visual areas in human cerebral cortex. *Proceedings of the National Academy of Sciences*, *93*, 2382–2386.

Dougherty, R. F., Koch, V. M., Brewer, A., Fischer, B., Modersitzki, J., & Wandell, B. (2003). Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *Journal of Vision*, *3*, 1–3.

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

Ferry, A. T., Öngür, D., An, X., & Price, J. L. (2000). Prefrontal cortical projections to the striatum in macaque monkeys: evidence for an organization related to prefrontal networks. *Journal of Comparative Neurology*, *425*, 447–470.

Fize, D., Vanduffel, W., Nelissen, K., Denys, K., Chef d'Hotel, C., Faugeras, O., & Orban, G. A. (2003). The retinotopic organization of primate dorsal V4 and surrounding areas: A functional magnetic resonance imaging study in awake monkeys. *Journal of Neuroscience*, *23*, 7395–7406.

Gattass, R., Sousa, A. P. B., & Gross, C. G. (1988). Visuotopic organization and extent of V3 and V4 of the macaque. *Journal of Neuroscience*, *8*, 1831–1845.

Girard, P., Lomber, S. G., & Bullier, J. (2002). Shape discrimination deficits during reversible deactivation of area V4 in the macaque monkey. *Cerebral Cortex*, *12*, 1146–1157.

Hadjikhani, N., Liu, A. K., Dale, A. M., Cavanagh, P., & Tootell, R. B. H. (1998). Retinotopy and color sensitivity in human visual cortical area V8. *Nature Neuroscience*, *1*, 235–241.

Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *Journal of Neuroscience*, *22*, 7195–7205.

Kelly, R. M., & Strick, P. L. (2003). Cerebellar loops with motor cortex and prefrontal cortex of a non-human primate. *Journal of Neuroscience*, *23*, 8432–8444.

Lewis, J. W., & Van Essen, D. C. (2000). Architectonic parcellation of parieto-occipital cortex and interconnected cortical regions in the Macaque monkey. *Journal of Comparative Neurology*, *428*, 79–111.

Lueck, C. J., Zeki, S., Friston, K. J., Deiber, M. P., Cope, P., Cunningham, V. J., Lammertsma, A. A., Kennard, C., & Frackowiak, R. S. J. (1989). The colour centre in the cerebral cortex of man. *Nature*, *340*, 386–389.

Lyon, D. C., & Kaas, J. H. (2002). Evidence for a modified v3 with dorsal and ventral halves in macaque monkeys. *Neuron*, *33*, 453–461.

McKeefry, D. J., & Zeki, S. (1997). The position and topography of the human colour centre as revealed by functional magnetic resonance imaging. *Brain*, *120*, 2229–2242.

Öngür, D., Ferry, A. T., & Price, J. L. (2003). Architectonic subdivision of the human orbital and medial prefrontal cortex. *Journal of Comparative Neurology*, *460*, 425–449.

Öngür, D., & Price, J. L. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cerebral Cortex*, *10*, 206–219.

Petrides, M., & Pandya, D. N. (2002). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *European Journal of Neuroscience*, *16*, 291–310.

Press, W. A., Brewer, A. A., Dougherty, R. F., Wade A. R., & Wandell, B. A. (2001). Visual areas and spatial summation in human visual cortex. *Vision Research*, *41*, 1321–1332.

Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*, 889–893.

Smith, A. T., Greenlee, M. W., Singh, K. D., Kraemer, F. M., & Hennig, J. (1998). The processing of first- and second-order motion in human visual cortex assessed by functional magentic resonance imaging (fMRI). *Journal of Neuroscience*, *18*, 3816–3830.

Tootell, R. B. H., & Hadjikhani, N. (2001). Where is "dorsal V4" in human visual cortex? Retinotopic, topographic and functional evidence. *Cerebral Cortex*, *11*, 298–311.

Tsao, D. Y., Vanduffel, W., Sasaki, Y., Fize, D., Knutsen, T. A., Mandeville, J. B., et al. (2003). Stereopsis activates V3A and caudal intraparietal areas in macaques and humans. *Neuron*, *39*, 555–568.

Van Essen, D. C. (2002). Surfacebased atlases of cerebellar cortex in human, macaque, and mouse. *Annals of the New York Academy of Sciences*, *978*, 468–479.

Van Essen, D. C. (2004). Organization of visual areas in macaque and human cerebral cortex. In L. Chalupa & J. S. Werner (Eds.), *The Visual Neurosciences* (pp. 507–521). Cambridge, MA: MIT Press.

Van Essen, D. C., Dickson, J., Harwell, J., Hanlon, D., Anderson, C. H., & Drury, H. A. (2001). An integrated software system for surface-based analyses of cerebral cortex. *Journal of American Medical Informatics Association*, *8*, 443–459.

Van Essen, D. C., Felleman, D. F., DeYoe, E. A., Olavarria, J. F., & Knierim, J. J. (1990). Modular and hierarchical organization of extrastriate visual cortex in the macaque monkey. *Cold Spring Harbor Symposia in Quantitative Biology*, *55*, 679–696.

Van Essen, D. C., Harwell, J., Hanlon, D., & Dickson, J. (2004). Surface-based atlases and a database of cortical structure and function. In S. H. Koslow & S. Subramaniam (Eds.), *Databasing the Brain: From Data to Knowledge (Neuroinformatics)*. New York: John Wiley & Sons, NJ (in press).

Van Essen, D. C., Maunsell, J. H. R., & Bixby, J. L. (1981). The middle temporal visual area in the macaque: Myeloarchitecture, connections, functional properties and topographic organization. *Journal of Comparative Neurology*, *199*, 293–326.

Van Essen, D. C., Newsome, W. T. N., & Bixby, J. L. (1982). The pattern of interhemispheric connections and its relationship to extrastriate visual areas in the macaque monkey. *Journal of Neuroscience*, *2*, 265–283.

Van Essen, D. C., Newsome, W. T., Maunsell, J. H. R., & Bixby, J. L. (1986). The projections from striate cortex (V1) to areas V2 and V3 in the macaque monkey: Asymmetries, areal boundaries, and patchy connections. *Journal of Comparative Neurology*, *244*, 451–480.

Van Essen, D. C., & Zeki, S. M. (1978). The topographic organization of rhesus monkey prestriate cortex. *Journal of Physiology*, *277*, 193–226.

Van Oostende, S., Sunaert, S., Van Hecke, P., Marchal, G., & Orban, G. A. (1997). The kinetic occipital (KO) region in man: An fMRI study. *Cerebral Cortex*, *7*, 690–701.

Wade, A. R., Brewer, A. A., Rieger, J. W., & Wandell, B. A. (2002). Functional measurements of human ventral occipital cortex: Retinotopy and color. *Philosophical Transactions of the Royal Society*, *357*, 963–973.

Walker, A. E. (1940). A cytoarchitectural study of the prefrontal areas of the macaque monkey. *Journal of Comparative Neurology*, *73*, 59–86.

Watson, J. D. G., Myers, R., Frackowiak, R. S. J., Hajnal, J. V., Woods, R. P., Mazziotta, J. C., et al. (1993). Area V5 of the human brain: Evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex*, *3*, 37–94.

2 Combined Human and Monkey fMRI Methods for the Study of Large-Scale Neuronal Networks in the Primate Brain

Zoe Kourtzi and Nikos K. Logothetis

The fundamental goal of studies in neuroscience is to understand the neuronal mechanisms that underlie primate behavior, at the level of both single neurons and neuronal ensembles. Substantial progress has been made in characterizing the response properties of single neurons involved in sensory, motor, and cognitive functions. In contrast, little is known about the collective properties of contiguous or distributed networks of neurons that underlie brain mechanisms. Functional magnetic resonance imaging (fMRI) provides global brain coverage and could therefore be used to investigate brain mechanisms at the level of distributed networks of neurons. Functional brain mapping is achieved by measuring local magnetic susceptibility alterations produced by changes in the concentration of deoxyhemoglobin in venous blood vessels. This blood-oxygenation-level-dependent (BOLD) contrast mechanism was successfully implemented in awake human subjects, in small animals, and recently in the nonhuman primate (for a review see Logothetis, 2002; 2003).

But how does the BOLD signal relate to the spike output which has been the primary neural measure thought to correlate with primate behavior? The localized increases in BOLD contrast have been assumed to reflect increases in neuronal activity. Intracortical recordings during fMRI in either anesthetized or alert monkeys have confirmed that the regional activations reflect local increases of neural responses. Recent combined fMRI and electrophysiological recordings (figure 2.1, plate 4) demonstrated that the BOLD signal primarily reflects the input of neuronal information and its processing at a given brain area rather than the output signals transmitted to other regions of the brain (Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001).

Here, we discuss how fMRI can be used in combination with other techniques to elucidate processing at the level of large-scale networks in the primate brain. Simultaneous imaging and electrophysiological recordings provide new insights into the mechanisms by which single neurons vs. neuronal ensembles contribute to the neural processing observed at a given cortical site. Furthermore, the strong correlation between the BOLD signal and the underlying neural responses emphasizes the importance of parallel fMRI studies in monkeys and humans investigating the same

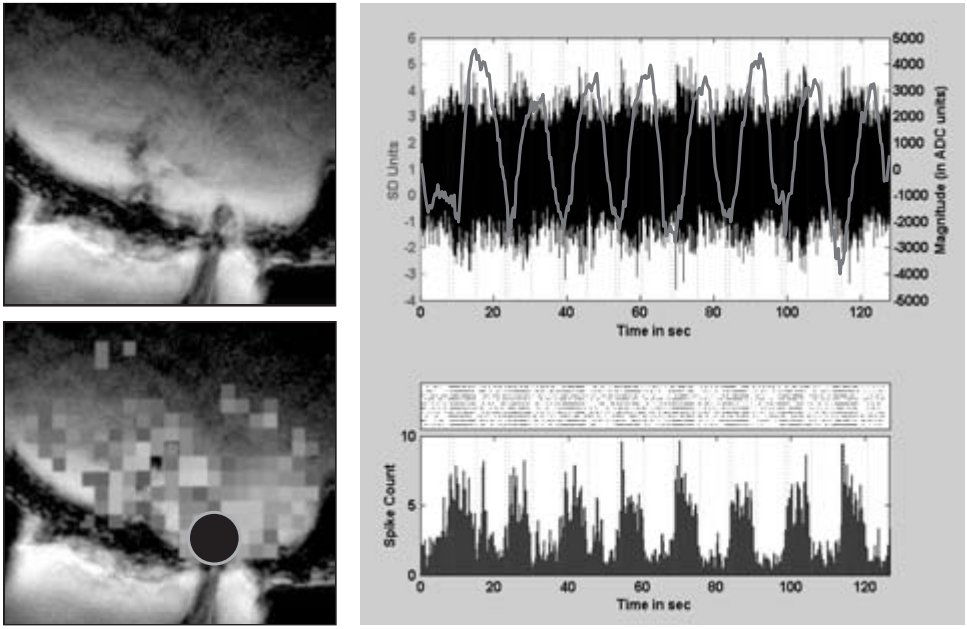


Figure 2.1
Combined neurophysiology and fMRI in alert, behaving monkeys. In the upper left panel an anatomical scan shows the position of the electrode tip. In the lower panel the activation elicited by a rotating checkerboard is superimposed on the anatomical scan. The top right panel shows the time course of the comprehensive signal (LFPs and spiking activity) together with the BOLD (red thick line) response. The bottom right illustrates the raster plots and peristimulus histograms of spiking activity. Each dot is an action potential and each bin shows the number of action potentials in 250 msec. See plate 4 for color version.

questions with similar tasks and stimuli. These studies are important for (1) bridging the gap between the extensive neurophysiological findings in monkeys and those reported in combined psychophysical and imaging investigations with humans, and (2) understanding similarities and/or differences in brain mechanisms that mediate behavior across species. Finally, the use of MRI-visible tracers and of electrical microstimulation combined with functional imaging are crucial for the study of connectivity in the primate brain.

Neural Correlates of the BOLD fMRI Signal

The firing rate of well-isolated neurons measured with microelectrodes has been the critical measure for comparing neural activity to sensory processing or behavior. Although a great deal has been learned from single-unit recordings, this technique

provides information mainly on single receptive fields but no access to subthreshold integrative processes or the associational operations taking place at a given site. Moreover, it suffers from an element of bias toward certain cell types (for references see Logothetis, 2002; 2003). That is, microelectrodes sample preferentially the somas or axons of large neurons that generate spikes that remain above the noise level over a greater distance from the cell than spikes from small neurons. Thus, it is possible that the measured spikes represent only very small populations of neurons, which in cortex are by and large the principal cells (e.g., pyramidal cells in cerebral cortex and Purkinje neurons in cerebellar cortex).

In contrast, functional MRI provides the possibility of testing large-scale neuronal networks but cannot directly measure neural responses. Instead it capitalizes on the interconnections among CBF, energy demand, and neural activity. Thus, comprehensive understanding of the relationship between the fMRI signal and the underlying neuronal activity is crucial for the interpretation of imaging findings in cognitive neuroscience. A number of studies in humans and animals have combined fMRI with electroencephalography (EEG; e.g., Menon et al., 1997; Krakow et al., 1999) or optical imaging recordings of intrinsic signals (Hess et al., 2000). However, optical imaging measures hemodynamic responses (Bonhoeffer & Grinvald, 1996) rather than directly neural activity, and EEG has poor spatial resolution and relatively imprecise localization of the electromagnetic field patterns associated with neural current flow.

Recently, studies using combined BOLD fMRI and intracortical recordings in anesthetized and conscious monkeys (figure 2.1) showed that the BOLD response directly reflects an increase in neural activity that correlates with electrical signals that are thought to represent synaptic inputs and local intracortical processing (Logothetis, 2002; 2003). In particular, these studies tested which cellular events contribute to the generation of the hemodynamic response measured in neuroimaging by examining the correlation of MUA, LFPs, and single neuron activity with the hemodynamic response (Logothetis et al., 2001). MUAs are a weighted sum of the extracellular action potentials of all neurons within a sphere of approximately 140–300 μm radius with the electrode at its center that are obtained approximately at a frequency band of 300–400 Hz (Grover & Buchwald, 1970; Legatt, Arezzo, and Vaughan, 1980; Henze et al., 2000). In contrast, LFPs are obtained at a frequency band less than 300 Hz and primarily represent slow events reflecting cooperative activity in neural populations. In all experiments, increases in the LFP range were greater in both spectral power and reliability. Furthermore, correlation analysis showed that LFPs are better predictors of the BOLD response than multiunit spiking (figure 2.2, plate 5). In fact, it was demonstrated that spike rate is nothing but a "fortuitous" predictor of the BOLD signal, simply because the firing of neurons itself usually happens to correlate with the LFPs. In cases in which there is a dissociation between these signals, BOLD is only predicted by the LFPs. For example, in sites exhibiting strong multiunit response adaptation MUA returned to the baseline approximately 2.5 seconds after stimulus onset, while

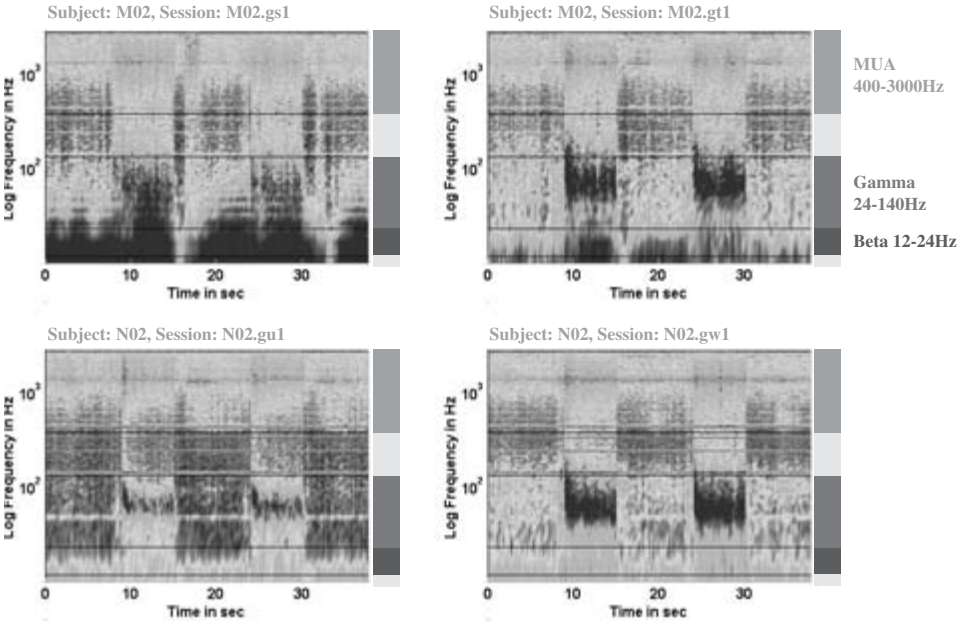


Figure 2.2
Time-dependent frequency analysis of the neural signals. Two sessions from two animals are shown (see labels). The spectrograms were computed for windows of 250 msec. Frequency is plotted in the y and time in the x axis. At the right of each plot the MUA and LFP ranges are indicated. The red (gamma band) and blue (beta band) sections indicated two of the usual EEG bands. Changes in the power of the signal in the gamma band was best correlated with the hemodynamic response. See plate 5 for color version.

the LFP signal remained elevated for the entire duration of the visual stimulus and was the only neural signal to be associated with the BOLD response.

Taken together, these results suggest that changes in the LFPs are more closely related to the evolution of the BOLD signal than changes in the spiking activity of single or multiple neurons. In other words, the BOLD signal mainly reflects the incoming specific or association inputs into an area and the processing of this input by the local cortical circuitry (including excitatory and inhibitory interneurons).

Comparison of fMRI and Physiological Findings

The well-established retinotopic organization of the visual system provides a benchmark for testing the specificity of the BOLD signal (e.g., figure 2.3, plate 6). Human fMRI studies have shown that retinotopy can be reliably demonstrated by using slowly

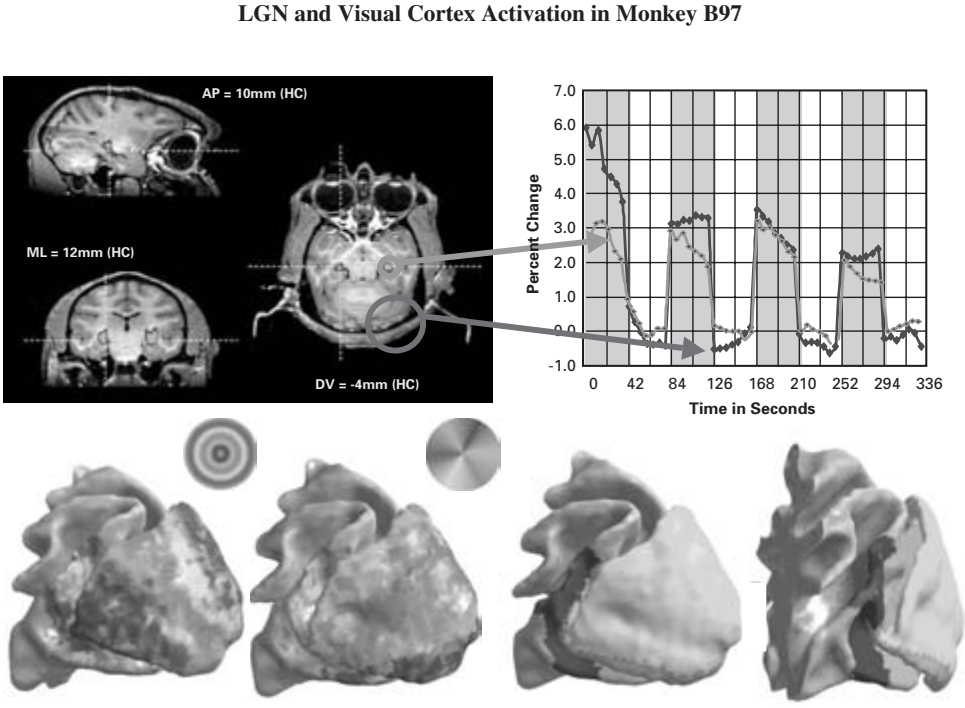


Figure 2.3

The upper left panels shows activation of lateral geniculate nucleus (LGN) and visual cortex in monkey B97. The yellow dotted lines indicate the LGN position. AP, anteroposterior; ML, mediolateral; DV, dorsoventral. The right panel shows the time course of the signal for the activated regions in LGN (red) and visual cortex (blue). The two lower left images show the retinotopic organization of the posterior visual areas as revealed with fMRI. Based on this organizations the boundaries of the areas can be defined (right two images); Cyan: V1, Magenta: V2, Yellow: V3, Red: V4, Blue: V4t, Green: MT(V5). See plate 6 for color version.

moving, phase-encoded retinotopic stimuli (Engel et al., 1994). When the same approach is used to study the retinotopical organization of the monkey visual areas (Brewer, Press, Logothetis, & Wandell, 2002), the maps obtained are in excellent agreement with those derived using anatomical and physiological techniques.

However, a number of recent studies on the functional specialization of visual areas suggest discrepancies between fMRI and physiological findings. For example, in a recent study Tolias and colleagues (Tolias et al., 2001) studied the brain areas processing motion information using an adaptation technique. Their results showed adaptation (i.e., decreased responses) for prolonged presentation of continuous motion in a single unchanging direction, but recovery form adaptation (i.e., rebound) for a

change in the direction of motion. The magnitude of this rebound was taken as an index of the average directional selectivity of neurons in any given activated area. The results confirmed previous electrophysiological studies revealing a distributed network of visual areas (V1, V2, V3, V5/MT) in the monkey that process information about the direction of visual motion. Surprisingly, however, strong activation was also observed in area V4, which is only weakly involved in motion processing (e.g., Desimone & Schein, 1987). Similarly, attentional effects in striate cortex have been very difficult to measure in monkey single cell electrophysiological recordings (Luck, Chelazzi, Hillyard, & Desimone, 1997; McAdams & Maunsell, 1999). Yet for similar tasks strong attentional effects are readily measurable with fMRI in human V1 (Gandhi, Heeger, & Boynton, 1999; Tong & Engel, 2000; Kastner & Ungerleider, 2000).

A possible explanation for such discrepancies is that in cases in which the activity of large projection neurons is shunted by concurrent modulatory input, the incoming afferent signals and the ongoing intracortical activity will elicit strong hemodynamic responses. In such cases spiking activity measured with microelectrodes will be a poor predictor of the BOLD response. As a result, the sensitivity of cortical areas may be influenced by "modulatory" input from other cortical areas that process different stimulus properties, which in and of itself is insufficient to drive the pyramidal cells recorded in a typical electrophysiology experiment. In such cases BOLD fMRI will reveal significant activation and will appear to provide results that do not match those of neurophysiology.

Beyond the Limited fMRI Resolution

Voxel Resolution

In animal experiments, very high resolution structural and functional imaging can be performed with small, tissue-compatible, intraosteally implantable radiofrequency coils (Logothetis, Merkle, Augath, Trinath, & Ugurbil, 2002). Tiny voxel sizes can be obtained with good signal and contrast to noise ratios revealing both structural and functional cortical architecture in great detail. Figure 2.4A (plate 7) shows an example of a T_2*-weighted Echo-Planar Image (EPI) obtained with an actual resolution of 125 $\times$ 125 μm^2 and a slice thickness of 720 μm. The contrast sensitivity of the image is sufficient to reveal the characteristic striation of the primary visual cortex. The dark line shown by the white arrow (Gen) is the well-known, approximately 200 μ thick Gennari line formed by the axons of pyramidal and spiny stellate cells contained in middle cortical layer (lamina IVB). Figure 2.4B shows fMRI correlation coefficient maps (in color) superimposed on the actual EPI (T_2*-weighted) images of a monkey during visual block-design stimulation. The sections are around the lunate sulcus, and activation extends into the primary and secondary visual cortices (V1 and V2). Both robust activation and good anatomical detail can be discerned.

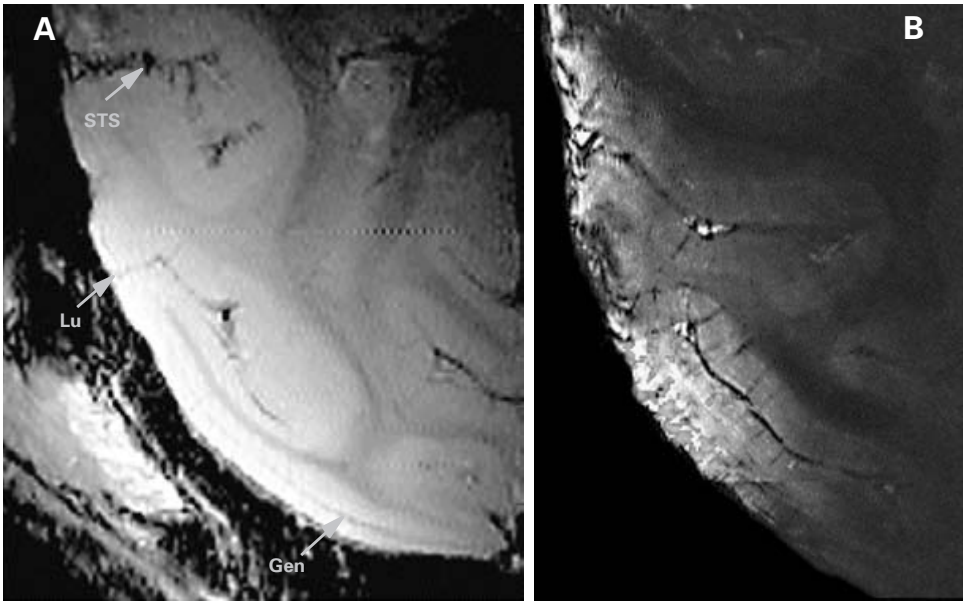


Figure 2.4
Anatomical and functional scans acquired with an implanted surface coil. (*A*) T2*-weighted EP image obtained with an actual resolution of $125 \times 125\,\mu m^2$ and a slice thickness of 720 μm. (*B*) Similar image from another animal with slightly different acquisition parameters. The resolution is sufficient to visualize the susceptibility effects produced by small cortical vessels with an average diameter of 120 μm. In color are the fMRI correlation coefficient maps. See plate 7 for color version.

fMRI Adaptation

One of the limitations of conventional fMRI paradigms that rely on the subtraction of activation between different stimulus types is that they average across neural populations that may respond homogeneously across stimulus changes or may be differentially tuned to different stimulus attributes. Thus, in most cases, it is impossible to infer the properties of the underlying imaged neural populations. A novel adaptation paradigm has been recently employed to study the properties of neuronal populations beyond this limited spatial resolution of fMRI. This paradigm capitalizes on the reduction of neural responses for stimuli that have been presented for prolonged time or repeatedly (Muller, Metha, Krauskopf, & Lennie, 1999; Lisberger and Movshon, 1999). A change in a specific stimulus dimension that elicits increased responses (i.e., rebound of activity) identifies neural populations that are tuned to the modified stimulus attributes (figure 2.5; plate 8).

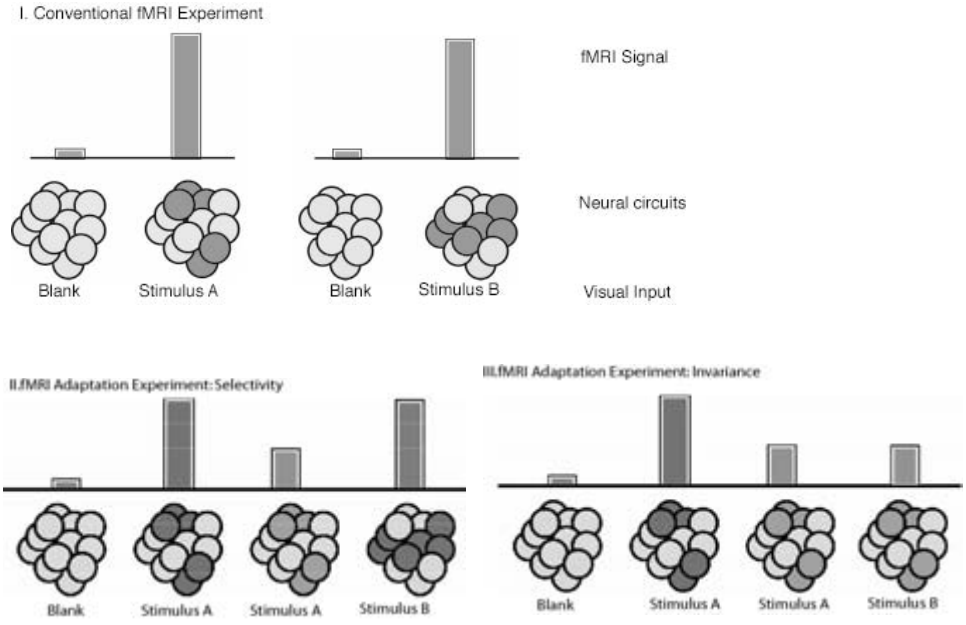


Figure 2.5
Conventional vs. adaptation fMRI paradigms. (*I*) Conventional imaging experiment: fMRI responses to two stimulus conditions (A and B) are compared to each other. If different neural subpopulations in the measured voxel encode the two stimuli, it is possible that the strength of the BOLD signal will be the same under these two conditions. Therefore, this conventional imaging experiment may fail to characterize the properties of these neural populations. (*II*) Adaptation experiment: stimulus A is shown for a prolonged time or repeatedly resulting in adaptation of the BOLD signal. If different neural subpopulations encode stimulus A and B then after presentation of stimulus B the signal shows a rebound; that is release from adaptation. (*III*) If the same neural subpopulations encode stimulus A and B, then the responses for stimulus B remain adapted after adaptation to stimulus A. See plate 8 for color version.

This paradigm has been used in both monkey and human fMRI studies as a sensitive tool that allows us to investigate (1) the selectivity of the neural populations and (2) the generalization of their responses within the imaged voxels. Specifically, recent imaging studies tested whether the neural populations in the early visual areas are tuned to visual features, e.g., orientation, color, direction of motion (Tootell et al., 1998; Engel & Furmanski, 2001; Huk and Heeger, 2001; Tolias et al., 2001). For example, after prolonged exposure to the adapting motion direction, observers were tested with the same stimulus in the same or in an orthogonal motion direction. Decreased fMRI responses were observed in MT when the stimuli were presented at the same motion direction as the adapting stimulus. However, recovery from this

adaptation effect was observed for stimuli presented at an orthogonal direction. These studies suggest that the neural populations in MT are tuned to direction of motion. Similarly, recent studies have shown stronger adaptation in MT/MST for coherently than for transparently moving plaid stimuli. These findings provide evidence that fMRI adaptation responses are linked to the activity of pattern-motion rather than component-motion cells in MT/MST (Huk & Heeger, 2002). Thus, these studies provide evidence that the fMRI signal can reveal neural selectivity consistent with the selectivity established by neurophysiological methods.

Recently, combined monkey (figure 2.6, plate 9) and human (figure 2.7) fMRI studies showed that coherent shape perception involves early (retinotopic) and higher (occipitotemporal) visual areas that may integrate local elements to global shapes at different spatial scales (Kourtzi, Tolias, Altmann, Augath, & Logothetis, 2003). fMRI responses across visual areas to collinear contours vs. random patterns were tested. The collinear patterns consisted of a number of similarly oriented elements embedded into a background of randomly oriented elements, while the random patterns consisted of a field of randomly oriented elements. Such displays yield the perception of a global figure in a randomly textured background and are thought to emerge from a segmentation process relying on the integration of the similarly oriented line-segments into global configurations (Field et al., 1993; Hess & Field, 1999 for review; Kovacs & Julesz, 1993; 1994). In the fMRI adaptation paradigm used, stimulus selectivity was deduced by changes in the course of adaptation of a pattern of randomly oriented elements. Accordingly, stronger increases of activity when local orientation changes in the adapting stimulus resulted in a collinear shape than a different random pattern were observed. In contrast to traditional approaches, selectivity for collinear shapes was shown not only in higher visual areas that are implicated in shape processing, but also in early visual areas where selectivity depended on the signal- (collinear elements) to-noise (random background elements) ratio within the receptive field size. These studies suggest that both early and higher visual areas are involved in the processing of global shapes at different spatial scales. Further human fMRI studies (Altmann, Bülthoff, & Kourtzi, 2003) showed decreased detection performance and fMRI activations when misalignment of the contour elements disturbed the perceptual coherence of the contours. However, grouping of the misaligned contour elements by disparity resulted in increased performance and fMRI activations, suggesting that similar neural mechanisms may underlie grouping of local elements to global shapes by different visual features (orientation or disparity). These studies provide additional evidence for the role of early perceptual organization processes and their interactions with higher stages of visual analysis in unified visual perception. Taken together, these findings provide evidence for common mechanisms in the human and nonhuman primate brain that are involved in coherent shape perception and bridge the gap between previous

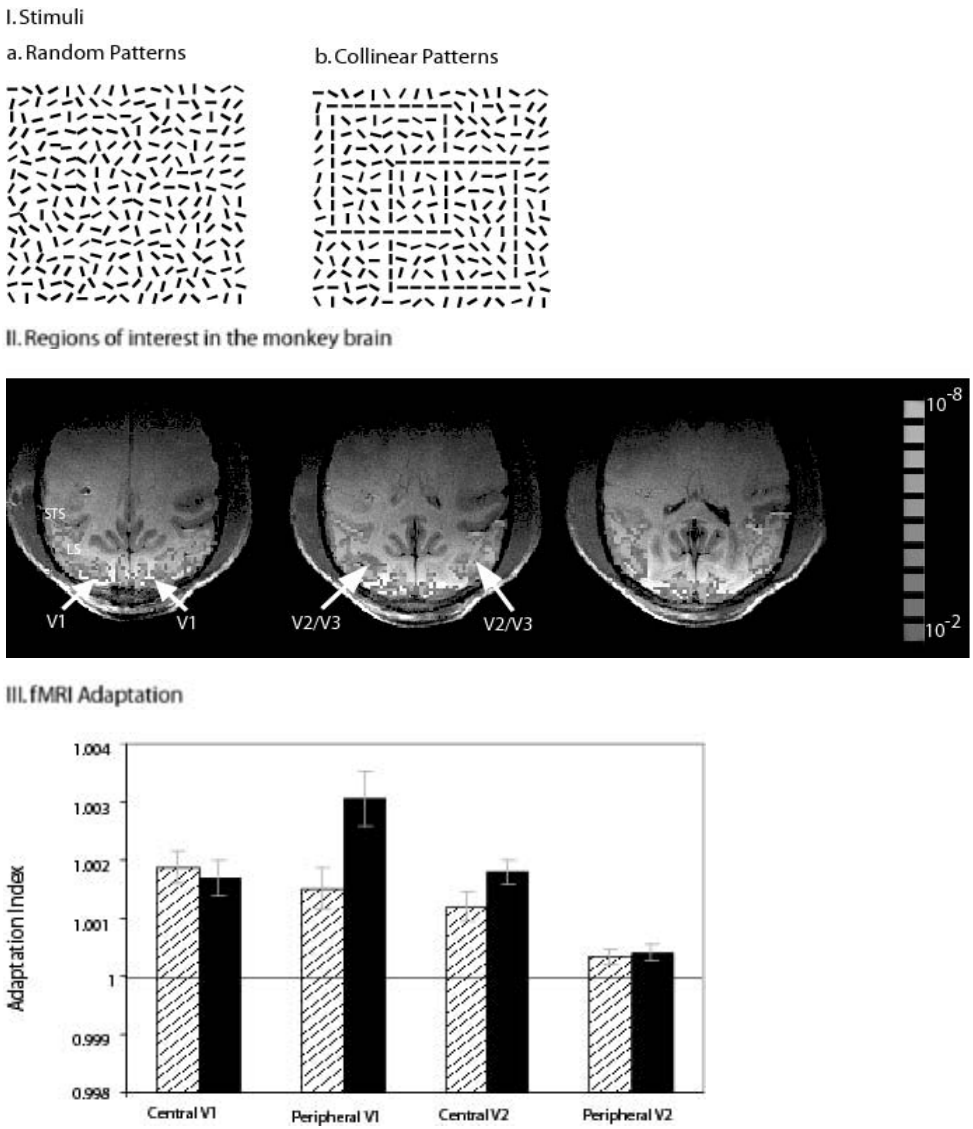


Figure 2.6

Monkey fMRI study on collinear shapes. (*I*) Stimuli rendered by oriented line segments: (*a*) random pattern used as the adapting stimulus and (*b*) collinear pattern used as the test stimulus. (*II*) Localization of the visual areas in the monkey brain. Three consecutive slices (posterior to anterior) from one subject showing the visual areas (V1, V2/V3) that were selected as regions of interest for the analysis of the adaptation experiment. These regions responded significantly more strongly to polar rotating rings than blank stimulation periods. Significance charts indicate

monkey electrophysiological and human fMRI findings on the neural processing of shapes.

Furthermore, recent human fMRI studies have used adaptation to test the selectivity and invariance of the responses of neural populations in the Lateral Occipital Complex (LOC), a region in the lateral occipital cortex extending anterior in the temporal cortex, that has been shown to be involved in shape processing (Kanwisher, Chun, McDermott, & Ledden, 1996; Malach et al., 1995). Adaptation across a change between two shapes provides evidence for a common neural representation invariant to that change, while recovery from adaptation suggests neural representations selective for specific shape properties. In particular, fMRI adaptation was used to test the effect of different stimulus transformations, namely position, size, orientation, and illumination change, on the BOLD signal in LOC responses (Grill-Spector et al., 1999; Grill-Spector & Malach, 2001). Adaptation was observed when the observers were presented repeatedly with identical images of objects. Stronger recovery from adaptation was shown across orientation or illumination changes compared to size and position changes. Interestingly, adaptation effects across orientation and size changes were observed more strongly in the anterior rather than the posterior regions of the LOC. Furthermore, fMRI adaptation was used to test whether the LOC is involved in the processing of object shape independent of low level image features that define the shape (figure 2.8, plate 10; Kourtzi & Kanwisher, 2001). An event-related fMRI

◀ **Figure 2.6**
the results of t-tests. The arrows point to the activated visual areas the borders of which were identified based on anatomical criteria. Major sulci: LS, lunate sulcus; STS, superior temporal sulcus. (*III*) fMRI adaptation results: We report responses during the adaptation experiment for areas V1 and V2/V3 where the activation was more robust. The weaker activation observed in area V4 was possibly due to the properties of the stimulus used for the adaptation experiment (statically flashed stimulus) that may not activate strongly these areas in the anesthetized monkey. We plot an fMRI adaptation index (fMRI responses in each condition / fMRI responses in the identical random pattern condition). A ratio of 1 (horizontal line) indicates adaptation. This adaptation index is plotted for the responses to the random-to-collinear pattern (solid bars) and to the different random pattern (striped bars) conditions across visual areas. The error bars indicate standard errors on the percent signal change averaged across scans and subjects. Collinearity effects were observed in peripheral V1 and central V2, but not in central V1, where only a small number of collinear elements was within the small size receptive fields, and peripheral V2, where the number of random background elements within the receptive field was possibly larger than the number of collinear elements. These fMRI adaptation results suggest that early visual areas contribute to the integration of local elements to global shapes based on the signal (collinear elements)-to-noise ratio (random background elements) within their receptive field. See plate 9 for color version.

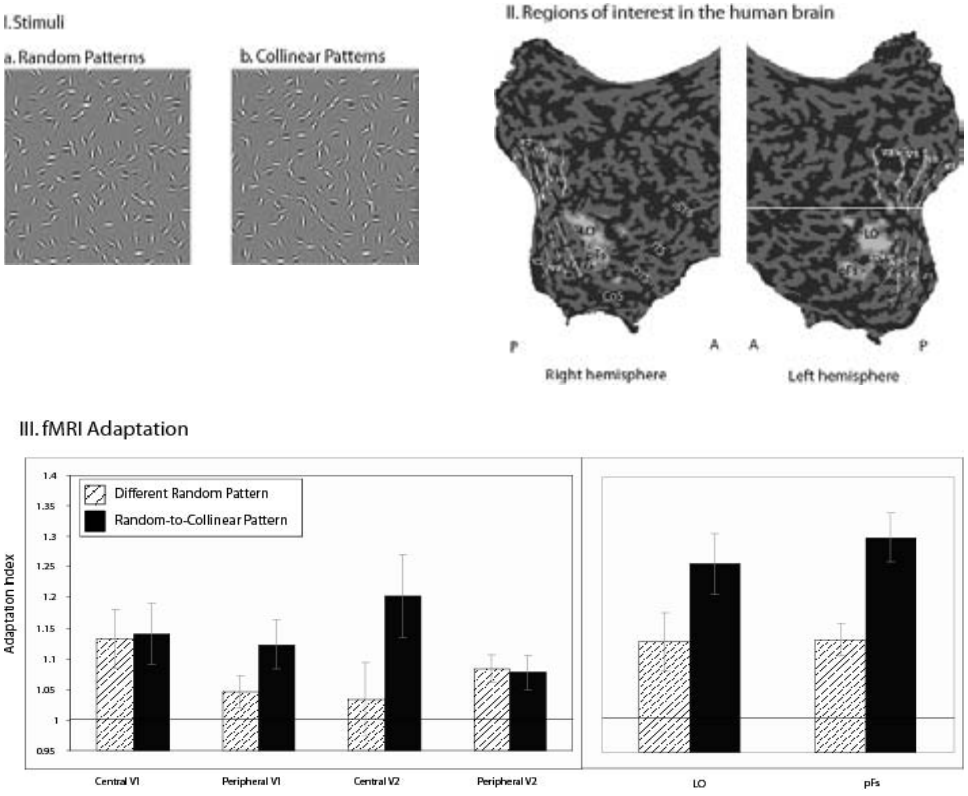


Figure 2.7

Human fMRI study on collinear shapes. (*I*) Stimuli rendered by Gabors. Examples of (*a*) the random patterns and (*b*) the collinear patterns used as stimuli. (*II*) Localization of the visual areas in the human brain. Functional activation maps for one subject showing the early retinotopic regions and the LOC. The functional activations are superimposed on flattened cortical surfaces of the right and left hemispheres. The sulci are coded in darker gray than the gyri and the antero-posterior orientation is noted by A and P. Major sulci are labeled: STS, superior temporal sulcus; ITS, inferior temporal sulcus; OTS, occipitotemporal sulcus; CoS, collateral sulcus. The borders (shown by lines) of the early visual regions (V1, V2, VP, V3, V3a, V4v) were defined with standard retinotopic techniques. The LOC was defined as the set of all contiguous voxels in the ventral occipitotemporal cortex that were activated more strongly ($p < 10^{-4}$) by intact than by scrambled images of objects. The posterior and (LO) and anterior regions (pFs) of the LOC were identified based on anatomical criteria. (*III*) fMRI Adaptation results: An fMRI adaptation index (percent signal change in each condition / percent signal change in the identical random pattern condition) reported for the random-to-collinear pattern (solid black bars) and the different random pattern (striped bars) conditions across visual areas. A ratio of 1 (horizontal line) indicates adaptation. This adaptation ratio is shown for central and peripheral subregions of V1 and

adaptation paradigm was employed, in which a pair of consecutively-presented stimuli was presented in each trial that lasted for 3 seconds. These studies showed adaptation in the LOC when the perceived shape was identical but the image contours differed (because occluding bars occurred in front of the shape in one stimulus and behind the shape in the other). In contrast, recovery from adaptation was observed when the contours were identical but the perceived shapes were different (because of a figure-ground reversal). Consistent with these results, adaptation was also shown for grayscale images and line drawings of the same objects (Kourtzi & Kanwisher, 2000) but not for object that differed in their 3D structure (i.e., convex vs. concave) (Kourtzi et al., 2003). These results suggest that neural populations in the LOC may not represent simple image features, such as contours, but higher-level shape information and 3D objects independent of image cues (i.e., shading and line contours).

Finally, an interesting aspect of adaptation is its relationship to learning and priming effects. Specifically, recent event-related fMRI studies have shown decreased activations in temporal and frontal areas for repeated presentation of objects (Buckner et al., 1998; James et al., 1999; Henson, Shallice, & Dolan, 2000; van Turennout, Ellmore, & Martin, 2000). It has been suggested that this adaptation effect is related to a psychophysical effect known as visual priming in which repeated presentation of a stimulus results in faster and more accurate observer performance in visual discrimination or object naming tasks (Schacter & Buckner, 1998; Wiggs & Martin, 1998; Naccache & Dehaene, 2001). Neurophysiological studies (Miller, Gochin, & Gross, 1991; Li, Miller, & Desimone, 1993) have also observed this repetition suppression effect and have proposed that it may reflect signals from neural populations that become smaller but more highly tuned to specific shape properties after the repeated presentations of objects (Desimone, 1996). As a result these neural populations become more selective to the repeated stimuli and may support more efficient behavioral responses.

In summary, adaptation is a powerful tool for studying the properties of networks of neurons in the human and nonhuman primate brain with imaging techniques beyond their limited spatial resolution. Although adaptation is a property of neural

◄ **Figure 2.7**

V2, posterior (LO) and anterior (pFs) subregions of the LOC. The error bars indicate standard errors on the percent signal change averaged across scans and subjects. Similar to the monkey fMRI adaptation study, collinearity effects were observed in peripheral V1 and central V2 consistent with the signal (collinear elements)-to-noise (random background elements) ratio within their receptive field. However, the collinearity effects in the LOC where the large receptive fields encode the whole stimulus that consisted of more background than collinear elements, suggest that neural populations in the LOC encode the perceived global shape rather than local configurations.

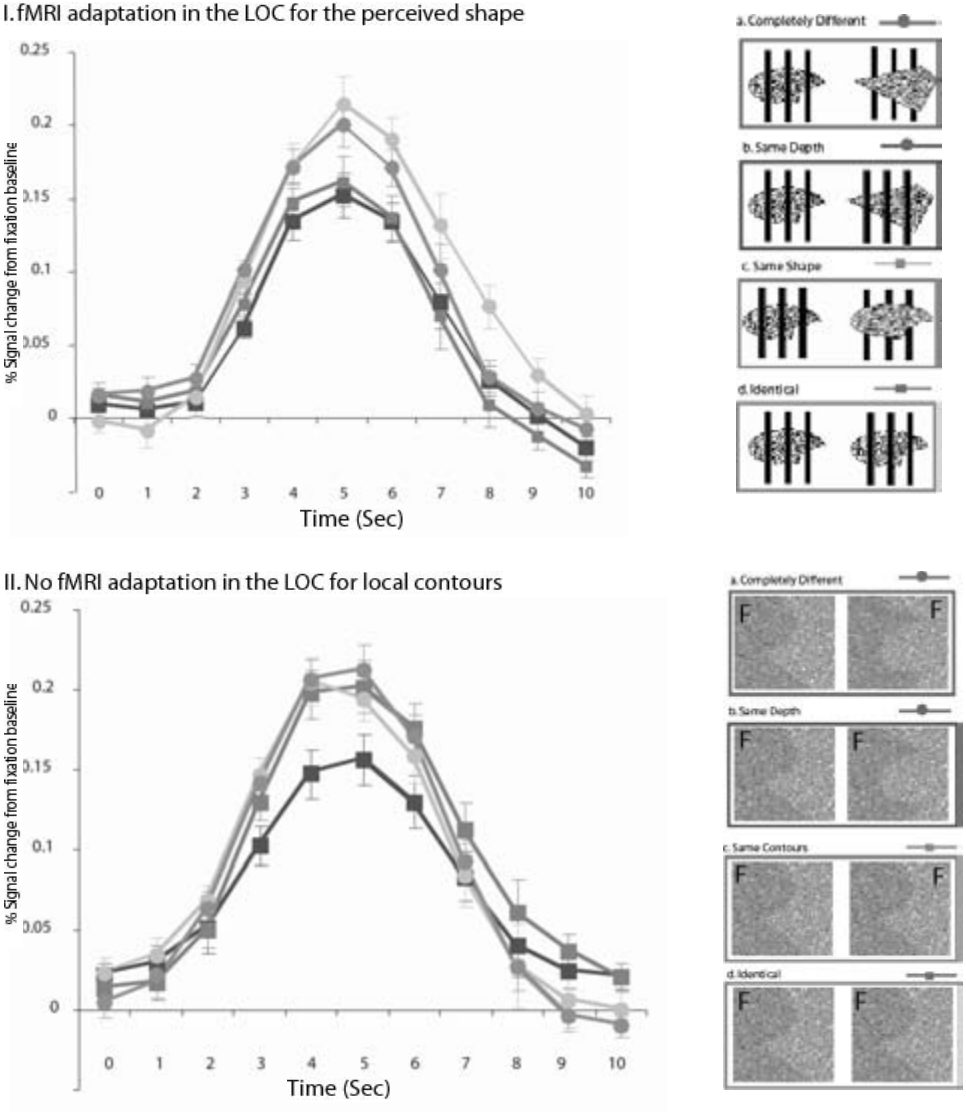


Figure 2.8

Shape processing in the human LOC. Data averaged across 10 subjects showing fMRI adaptation effects in the LOC, that is decreased responses (% signal change from fixation baseline) for identical images of objects (compared to the responses for different objects in a trial. (*I*) Adaptation is shown for images that have the same perceived shape but different contours due to occlusion. (*II*) In contrast, no adaptation is shown for images that when rendered stereoscopically have the same contours but different perceived shape due to figure ground reversal (F indicates the shape perceived as the figure in front of the background for each image). These fMRI adaptation results suggest that neural populations in the LOC encode the perceived shape of objects rather than their local contours. See plate 10 for color version.

responses, the relationship between the adaptation of the BOLD signal and neuronal activity is currently not known. Simultaneous recordings of the BOLD signal and electrophysiological activity during adaptation are likely to provide further insights about the relationship between BOLD and neuronal adaptation.

Connectivity Studies with MRI

MRI with Paramagnetic Tracers

The topographical connectivity between different brain areas has been primarily examined by degeneration (Jones & Powell, 1970; Seltzer & Pandya, 1978) and anterograde and retrograde tracer techniques (Saint-Cyr, Ungerleider, & Desimone, 1990; Felleman & Van Essen, 1991; Saleem, Suzuki, Tanaka, & Hashikawa, 2000). One limitation of these studies is that they require fixed processed tissue and therefore cannot be used in longitudinal studies examining an entire circuit in the same subjects.

However, MRI visible tracers that are infused into a specific brain region and are transported anterogradely or retrogradely along the axon may enable us to study connectivity in the living animal. For example, a recent study using manganese (Mn^{2+}), an MRI-visible contrast agent, provided a detailed account of both the specificity and the transsynaptic transfer of this substance. Injections were made in the striatum (Saleem et al., 2002). Its projections were confirmed histologically in the same animals by injecting WGA-HRP at the same sites where $MnCl_2$ had been injected. The size and location of the projection foci in the striatal targets were comparable to those found in both the MR and histology images. By injecting WGA-HRP at the same sites as $MnCl_2$, we also confirmed for each animal the absence of a direct connection from the injection sites to various brain structures (e.g., thalamic nuclei). In this study, manganese was actually found in a number of structures receiving no direct projections from the injected sites. Such paramagnetic tracer studies may be used to validate and further develop noninvasive fiber tracking techniques such as diffusion tensor MRI (see, for example, LeBihan et al., 2001) that permit the study of connectivity even in the human brain.

MR Imaging and Electrical Microstimulation

Electrical microstimulation is established as an important tool for the study of areal representations and the functional inta-areal connectivity. Combination of fMRI with electrical microstimulation could enhance our knowledge of functional connectivity and cortical organization in the primate brain. A new method was recently developed for combined fMRI and microstimulation studies by using specially constructed microelectrodes to stimulate directly a selected subcortical or cortical area while simultaneously measuring changes in BOLD signal (Logothetis et al., 2001). The exact location of the stimulation site was determined by means of anatomical scans as well as by the

study of the physiological properties of neurons. Electrical stimulation was delivered using a biphasic pulse generator attached to a constant-current stimulus isolation unit. Local microstimulation of striate cortex yielded both local BOLD signals and activation of areas V2, V3, and MT. Microstimulation of dLGN resulted in the activation of striate cortex as well as areas V2, V3, and MT. The findings show that microstimulation combined with fMRI can be an exquisite tool for finding target areas of electrophysiological interest and studying their functional connectivity.

Conclusions

The suitability of MRI for functional brain mapping in awake human subjects as well as in animals such as rats, cats, and monkeys has been firmly established. MRI studies in high magnetic fields, in which voxels may contain as few as 600–800 cortical neurons, can help us understand how neural networks are organized, and how small cell assemblies contribute to the activation patterns revealed in fMRI. The combination of this technique with electrophysiology has confirmed that the areal activations measured in MR neuroimaging do indeed reflect local increases in neural activity. In addition, it has been demonstrated that fMRI responses mostly reflect the input of a given cortical area and its local intracortical processing. This strong correlation between the BOLD signal and the underlying neural responses emphasizes the importance of combined fMRI studies in monkeys and humans in understanding the neural mechanisms that mediate behavior across species. Further simultaneous imaging and electrophysiological recordings are needed for investigating the mechanisms by which single neurons vs. neuronal ensembles contribute to the neural code that underlies behavior. Finally, the combination of fMRI with tracer and microstimulation techniques appears to be ideal for the study of connectivity in the living animal and for the validation and advancement of noninvasive techniques that can be used for the study of the human brain.

References

Altmann, C. F., Bülthoff, H. H., & Kourtzi, Z. (2003). Perceptual organization of local elements into global shapes in the human visual cortex. *Current Biology*, *13*(4), 342–349.

Bonhoeffer, T., & Grinvald, A. (1996). Optical Imging Based on Intrinsic Signals. In A. W. Toga & J. C. Mazziotta (Eds.), *Brain Mapping, The Methods* (pp. 55–97). New York: Academic Press.

Brewer, A. A., Press, W. A., Logothetis, N. K., & Wandell, B. A. (2002). Visual areas in macaque cortex measured using functional magnetic resonance imaging. *Journal of Neuroscience*, *22*(23), 10416–10426.

Buckner, R. L., Goodman, J., Burock, M., Rotte, M., Koutstaal, W., Schacter, D., et al. (1998). Functional-anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI. *Neuron*, *20*(2), 285–296.

Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences*, *93*(24), 13494–13499.

Desimone, R., & Schein, S. J. (1987). Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *Journal of Neurophysiology*, *57*(3), 835–868.

Engel, S. A., & Furmanski, C. S. (2001). Selective adaptation to color contrast in human primary visual cortex. *Journal of Neuroscience*, *21*(11), 3949–3954.

Engel, S. A., Rumelhart, D. E., Wandell, B. A., Lee, A. T., Glover, G. H., Chichilnisky, E. J., et al. (1994). fMRI of human visual cortex. *Nature*, *369*(6481), 525.

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

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local "association field." *Vision Research*, *33*(2), 173–193.

Gandhi, S. P., Heeger, D. J., & Boynton, G. M. (1999). Spatial attention affects brain activity in human primary visual cortex. *Proceedings of the National Academy of Sciences*, *96*(6), 3314–3319.

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, *24*(1), 187–203.

Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psycholigica (Amsterdam)*, *107*(1–3), 293–321.

Grover, F. S., & Buchwald, J. S. (1970). Correlation of cell size with amplitude of background fast activity in specific brain nuclei. *Journal of Neurophysiology*, *33*(1), 160–171.

Henson, R., Shallice, T., & Dolan, R. (2000). Neuroimaging evidence for dissociable forms of repetition priming. *Science*, *287*(5456), 1269–1272.

Henze, D. A., Borhegyi, Z., Csicsvari, J., Mamiya, A., Harris, K. D., & Buzsaki, G. (2000). Intracellular features predicted by extracellular recordings in the hippocampus in vivo. *Journal of Neurophysiology*, *84*(1), 390–400.

Hess, R., & Field, D. (1999). Integration of contours: new insights. *Trends in Cognitive Science*, *3*(12), 480–486.

Hess, A., Stiller, D., Kaulisch, T., Heil, P., & Scheich, H. (2000). New insights into the hemodynamic blood oxygenation level-dependent response through combination of functional magnetic resonance imaging and optical recording in gerbil barrel cortex. *Journal of Neuroscience*, *20*(9), 3328–3338.

Huk, A. C., & Heeger, D. J. (2002). Pattern-motion responses in human visual cortex. *Nature Neuroscience*, *5*(1), 72–75.

Huk, A. C., Ress, D., & Heeger, D. J. (2001). Neuronal basis of the motion aftereffect reconsidered. *Neuron*, *32*(1), 161–172.

James, T. W., Humphrey, G. K., Gati, J. S., Menon, R. S., & Goodale, M. A. (1999). Repetition priming and the time course of object recognition: An fMRI study. *NeuroReport*, *10*(5), 1019–1023.

Jones, E. G., & Powell, T. P. (1970). An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. *Brain*, *93*(4), 793–820.

Kanwisher, N., Chun, M. M., McDermott, J., & Ledden, P. J. (1996). Functional imagining of human visual recognition. *Cognitive Brain Research*, *5*(1–2), 55–67.

Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annual Reviews in Neuroscience*, *23*, 315–341.

Kourtzi, Z., Erb, M., Grodd, W., & Bulthoff, H. H. (2003). Representation of the perceived 3-D object shape in the human lateral occipital complex. *Cerebral Cortex*, *13*(9), 911–920.

Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *Journal of Neuroscience*, *20*(9), 3310–3318.

Kourtzi, Z., & Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science*, *293*(5534), 1506–1509.

Kourtzi, Z., Tolias, A. S., Altmann, C. F., Augath, M., & Logothetis, N. K. (2003). Integration of local features into global shapes: monkey and human FMRI studies. *Neuron*, *37*(2), 333–346.

Kovacs, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one: effect of closure in figure-ground segmentation. *Proceedings of the National Academy of Sciences*, *90*(16), 7495–7497.

Kovacs, I., & Julesz, B. (1994). Perceptual sensitivity maps within globally defined visual shapes. *Nature*, *370*(6491), 644–646.

Krakow, K., Woermann, F. G., Symms, M. R., Allen, P. J., Lemieux, L., Barker, G. J., et al. (1999). EEG-triggered functional MRI of interical epileptiform activity in patients with partial seizures. *Brain*, *122*, 1679–1688.

Le Bihan, D., Mangin, J. F., Poupon, C., Clark, C. A., Pappata, S., Molko, N., et al. (2001). Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging*, *13*(4), 534–546.

Legatt, A. D., Arezzo, J., & Vaughan, H. G. (1980). Averaged multiple unit activity as an estimate of phasic changes in local neuronal activity: effects of volume-conducted potentials. *Journal of Neuroscientific Methods*, *2*(2), 203–217.

Li, L., Miller, E. K., & Desimone, R. (1993). The representation of stimulus familiarity in anterior inferior temporal cortex. *Journal of Neurophysiology*, *69*(6), 1918–1929.

Lisberger, S. G., & Movshon, J. A. (1999). Visual motion analysis for pursuit eye movements in area MT of macaque monkeys. *Journal of Neuroscience*, *19*(6), 2224–2246.

Logothetis, N., Merkle, H., Augath, M., Trinath, T., & Ugurbil, K. (2002). Ultra high-resolution fMRI in monkeys with implanted RF coils. *Neuron*, *35*(2), 227–242.

Logothetis, N. K. (2002). The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. *Philosophical Transactions of the Royal Society London B, Biological Sciences*, *357*(1424), 1003–1037.

Logothetis, N. K. (2003). The underpinnings of the BOLD functional magnetic resonance imaging signal. *Journal of Neuroscience*, *23*(10), 3963–3971.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150–157.

Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology*, *77*(1), 24–42.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences*, *92*(18), 8135–8139.

McAdams, C. J., & Maunsell, J. H. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *Journal of Neuroscience*, *19*(1), 431–441.

Menon, V., Ford, J. M., Lim, K. O., Glover, G. H., & Pfefferbaum, A. (1997). Combined event-related fMRI and EEG evidence for temporal-parietal cortex activation during target detection. *Neuroreport*, *8*, 3029–3037.

Miller, E. K., Gochin, P. M., & Gross, C. G. (1991). Habituation-like decrease in the responses of neurons in inferior temporal cortex of the macaque. *Visual Neuroscience*, *7*(4), 357–362.

Muller, J. R., Metha, A. B., Krauskopf, J., & Lennie, P. (1999). Rapid adaptation in visual cortex to the structure of images. *Science*, *285*(5432), 1405–1408.

Naccache, L., & Dehaene, S. (2001). The priming method: Imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cerebral Cortex*, *11*(10), 966–974.

Saint-Cyr, J. A., Ungerleider, L. G., & Desimone, R. (1990). Organization of visual cortical inputs to the striatum and subsequent outputs to the pallido-nigral complex in the monkey. *Journal of Computational Neurology*, *298*(2), 129–156.

Saleem, K. S., Pauls, J. M., Augath, M., Trinath, T., Prause, B. A., Hashikawa, T., et al. (2002). Magnetic resonance imaging of neuronal connections in the macaque monkey. *Neuron*, *34*(5), 685–700.

Saleem, K. S., Suzuki, W., Tanaka, K., & Hashikawa, T. (2000). Connections between anterior inferotemporal cortex and superior temporal sulcus regions in the macaque monkey. *Journal of Neuroscience*, *20*(13), 5083–5101.

Schacter, D. L., & Buckner, R. L. (1998). Priming and the brain. *Neuron*, *20*(2), 185–195.

Seltzer, B., & Pandya, D. N. (1978). Afferent cortical connections and architectonics of the superior temporal sulcus and surrounding cortex in the rhesus monkey. *Brain Research*, *149*(1), 1–24.

Tolias, A. S., Smirnakis, S. M., Augath, M. A., Trinath, T., & Logothetis, N. K. (2001). Motion processing in the macaque: revisited with functional magnetic resonance imaging. *Journal of Neuroscience*, *21*(21), 8594–8601.

Tong, F., & Engel, S. A. (2001). Interocular rivalry revealed in the human cortical blind-spot representation. *Nature*, *411*(6834), 195–199.

Tootell, R. B., Hadjikhani, N. K., Vanduffel, W., Liu, A. K., Mendola, J. D., Sereno, M. I., et al. (1998). Functional analysis of primary visual cortex (V1) in humans. *Proceedings of the National Academy of Sciences*, *95*(3), 811–817.

Van Turennout, M., Ellmore, T., & Martin, A. (2000). Long-lasting cortical plasticity in the object naming system. *Nature Neuroscience*, *3*(12), 1329–1334.

Wiggs, C. L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinions in Neurobiology*, *8*(2), 227–233.

3 Evolution of the Human Brain and Comparative Cyto- and Receptor Architecture

Karl Zilles

Many morphological features of the human brain are similar to those of other primate brains, particularly ape brains (e.g., orangutan; see Zilles & Rehkämper, 1988). The human brain, however, is more than simply a large monkey or ape brain. It differs from other primate or mammalian brains by the most intensive growth of the evolutionary youngest parts of the brain, i.e., the neocortex with its primary and secondary sensory regions, and particularly by the exceptional size of the multimodal cortical regions ("association cortex") as well as the white matter in the forebrain (figure 3.1). This disproportional enlargement of multimodal regions is a structural indicator of a corresponding functional differentiation of increasingly complex neural mechanisms during human brain evolution. An extensive rewiring and increase of synaptic connections as well as changes in system functions can be hypothesized. Even some evolutionary relatively conservative brain regions (e.g., hippocampus) show such changes during primate brain development.

The disproportional increase of specific brain regions can best be demonstrated by comparing the sizes of different forebrain compartments on an allometric scale (Stephan, 1975). This comparison takes as reference the group of *Tenrecinae* (basal insectivores) which shows the most "primitive" features of brain morphology. The slope α of the regression of brain to body size is very similar amongst nearly related groups of mammals. It amounts to $\alpha = 0.65$ in basal insectivores, if brain and body weights are plotted in a double logarithmic scale (Stephan, 1975). The differential evolutionary progression of the neocortex, archicortex (hippocampus, entorhinal cortex and neighboring allocortical areas), and paleocortex (olfactory bulb and olfactory cortical regions), as well as the white matter of the forebrain, was measured in insectivores, prosimians, simians, and *Homo sapiens* (Stephan, Baron, & Frahm, 1991). The progression index indicates how many times larger an actual brain structure is in a given species or a group of nearly related species if compared with the size of this brain structure in a hypothetical member of the *Tenrecinae* of equal body size. The size of the brain of the hypothetical *Tenrecinae* species or its parts can be estimated by an allometric upscaling of the small-sized "archetypical" members of the Tenrecinae to

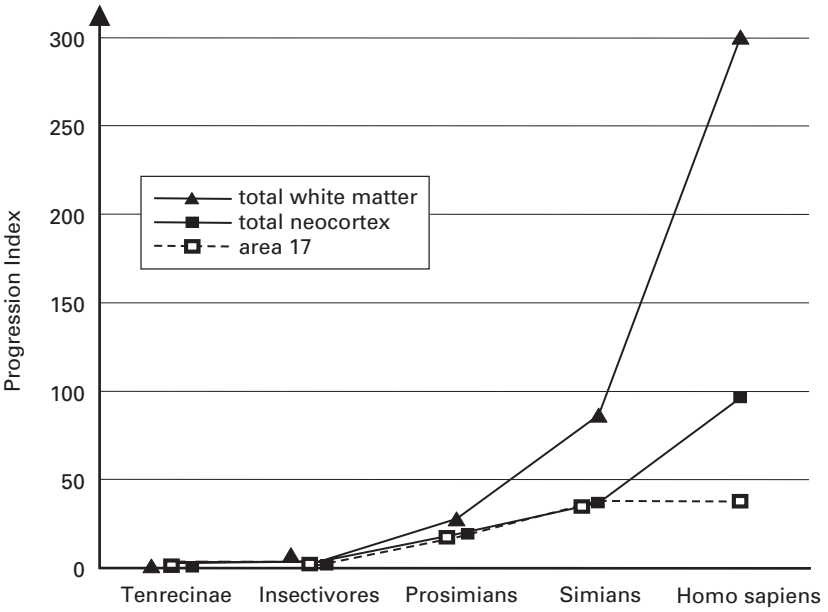


Figure 3.1
Progression indices of neocortex, white matter of the forebrain and primary visual cortex (area 17 of Brodmann [1909] including the underlying white matter) in *Tenrecinae* (Madagascar hedgehogs as representatives of the most basal mammals), insectivores, prosimians, simians and humans. For explanation of progression index, see text. (Data from Stephan et al., 1991.)

the body size of the species under investigation, e.g., *Homo sapiens*. This type of analysis demonstrates that the volume of the paleocortex (rhinencephalon) does not increase between *Tenrecinae* and prosimians, Old and New World monkeys and humans (not shown in figure 3.1). The neocortex and particularly the white matter of the forebrain increase, however, to an exceptionally high degree (figure 3.1). The allometric volume increase of neocortex versus area 17 (primary visual neocortical area) indicates that other areas than primary sensory regions, i.e., the multimodal association cortex, contribute most to the impressive growth of the neocortex in the human brain.

Evolution of Hemispheric Shape and Cortical Folding in Primates

The increase in neocortical volume is associated with an increase of the number of cortical minicolumns (Armstrong et al., 1991). This fundamental condition of brain growth results in a considerable enlargement of the cortical surface. Since the volume of the brain is constrained, among other things, by the relation between the diame-

ter of the birth canal in females and the diameter of the head of the newborn as well as by the need for minimizing the lengths of fiber connections between different cortical regions, cortical surface folding (gyrification) (Armstrong et al., 1991) and changes in the overall shape of the brain appear to be an efficient evolutionary solution of this problem.

The change in hemispheric shape can be demonstrated by a novel morphing procedure (elastic transformation, "warping"; Crivello et al., 2002; Mohlberg et al., 2002; Schormann & Zilles, 1998), which transforms the volume representation of an actual brain or endocast into that of a reference brain or endocast that can be freely defined. Since this procedure is completely algorithmically driven, it allows (by calculation of the deformation fields after warping) the quantitative analysis of evolutionary changes in hemispheric shape between nonhuman primate and human brains as well as between endocasts of hominids and *Homo sapiens*. The brains or endocasts are virtual voxel-based representations acquired by magnetic resonance imaging and 3D computer reconstructions. The deformation fields reveal the mosaic-like evolution of the neocortex from nonhuman primates to *Homo sapiens*, particularly of the orbital prefrontal cortex and parieto-temporo-occipital association cortex (figure 3.2, plate 11).

As an example, figure 3.2A shows changes in overall brain shape, when a virtual bonobo brain is warped to a virtual human brain after volume normalization of both brains. The most pronounced changes in brain shape are found in the orbito-prefrontal cortical region (red arrows in figure 3.2A). This region of the association cortex is found to increase and protrude when the cortical surface of a volume-normalized bonobo brain is fitted to that of a brain of *Homo sapiens*. Figure 3.2B shows the changes in shape, when an endocast of *Paranthropus boisei* is warped to a mean human endocast (averaged over 11 specimens) after volume normalization of both endocasts. The strongest deformations—either relative local increase (bulbing out, orange to yellow-coded) or decrease (compression, blue to white-coded) of the surface of the *Paranthropus* endocast after warping to the *Homo sapiens* endocast—are found over the orbito-prefrontal cortex (red arrows in figure 3.2B), the lateral sensorimotor to lateral prefrontal cortices of the right hemisphere (small arrows in figure 3.2B)), and the temporo-occipital regions of both hemispheres (small arrows in figure 3.2B). These findings clearly indicate that the human brain is not only larger than the bonobo or *Paranthropus* brains, but that the absolute increase has also led to localized remodeling of hemispheric shape, which may be an indicator of anatomically as well as functionally specific reorganizations of the human neocortex.

Absolute increase in cortical volume leads not only to changes in hemispheric shape, but also to folding of the cortical surface (gyrification). The changes in gyrification during primate brain evolution has been repeatedly demonstrated by us (Armstrong et al., 1991; Armstrong, Zilles, Curtis, & Schleicher, 1991; Armstrong, Zilles, &

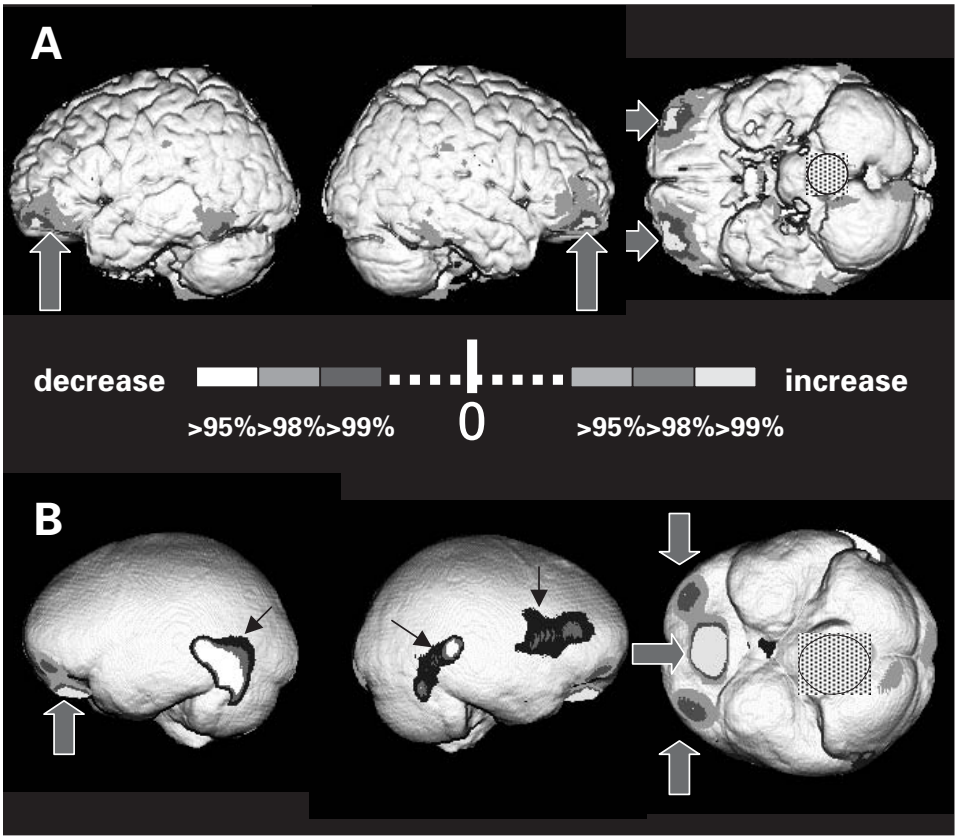


Figure 3.2
Changes in shape (*A*) between a bonobo brain and a brain of *Homo sapiens*, and (*B*) between an endocast of *Paranthropus boisei* (KNM-ER 732) and a mean endocast (averaged over 11 specimens) of *Homo sapiens*. The evolutionary changes in hemispheric shape are calculated as deformation fields after the volumes of all species have been normalized (Zilles et al., 2001). All voxels of each brain or endocast were algorithmically deformed in x-, y-, and z-directions, and the deformation vector of each voxels was registered. All deformation vectors represent the deformation field after warping. Only voxels with deformation vectors passing the right (indicates local bulging out of the surface) or left (indicates local compression of the surface) 95, 98, or 99 percent levels of the frequency distribution of all deformation vectors of a specimen are color coded. Zero indicates no changes between species. Dark blue (95 percent) to light blue (98 percent) to white (≥99 percent) regions indicate the highest local compressions (*Homo* < *Paranthropus*) when the surfaces of the brain of bonobo or the endocast of *Paranthropus* endocast are fitted to that of the brain or endocast of *Homo sapiens*. Orange (95 percent) to red (98 percent) to yellow (≥99 percent) indicate local bulging out (*Homo* > *Paranthropus*) of the surface in the respective warps. See plate 11 for color version.

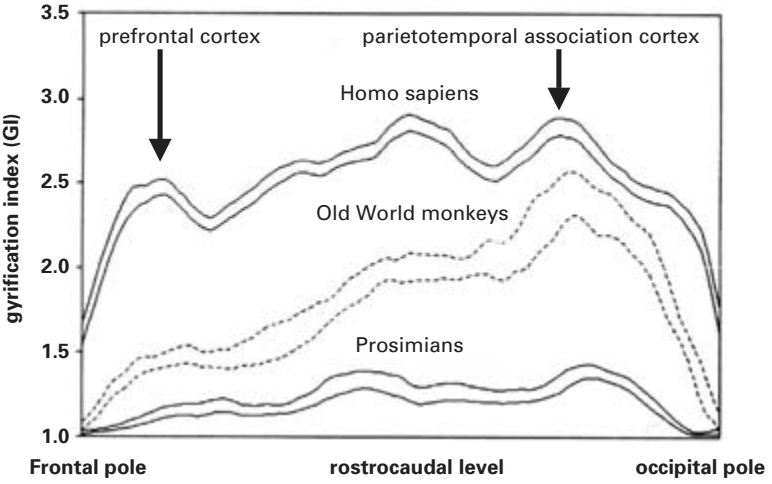


Figure 3.3
Gyrification index GI registered in a rostrocaudal sequence in brains of 29 different primate species. The primate species are classified as belonging to prosimians (20 different species), Old World monkeys (8 different species) or *Homo sapiens* (61 individuals). The double curves indicate the 95 percent confidence limits in each group. (Modified after Zilles et al., 1988.)

Schleicher 1993; Armstrong, Zilles, Omran, & Schleicher, 1995; Zilles, Armstrong, Schleicher, & Kretschmann, 1988; Zilles, Armstrong, Moser, Schleicher, & Stephan 1989). A quantitative measure of the intensity of gyrification is the gyrification index GI. This index is the ratio between the total length of the outer cortical contour in serial sections through the brain (superficially exposed cortical surface plus cortical surface hidden in the sulci) and the contour length of only the superficially exposed cortical surface in the same serial sections. The mean GI of the brain of *Homo sapiens* indicates that approximately two-thirds of the cortical surface is buried in the sulci, whereas the GI in Old World monkeys is clearly lower, and more so in prosimians. A comparative analysis of 29 different primate species shows not only the mean GIs (averaged over the whole brain) of different primates, but also localized changes, when the GI is sequentially registered throughout each brain from the frontal to the occipital poles (figure 3.3).

These GI-curves (figure 3.3) reach local maxima in prosimians, Old World monkeys, and *Homo sapiens* in the parietotemporal association cortex. A pronounced second local maximum is found at the level of the prefrontal region mainly in humans. A third local maximum is most clearly expressed in the human brain at the level of the posterior temporal and anterior occipital association cortices. These results clearly demonstrate that the most intensive cortical folding takes place in the human brain,

and exceeds that of Old World monkeys particularly in the prefrontal and parieto-temporo-occipital association cortices. In conclusion, the human neocortex has increased particularly in its association regions. The GI and the changes in hemispheric shape reflect accordingly this regional-specific evolution of the human forebrain.

Comparative Cytoarchitecture of the Primate Cerebral Cortex

On the basis of its common six-layered structure, the cytoarchitecture of the neocortex in prosimians, New World and Old World monkeys as well as humans reveals considerable adaptations within and between the different species to the functional requirements. Only a few examples will be presented in this chapter.

The motor cortex of New World monkeys (figure 3.4B), Old World monkeys (figure 3.4C) and *Homo sapiens* (figure 3.4D) is characterized by a specific variation of its basic bauplan, i.e., the loss of a well recognizable inner granular layer (layer IV) before and shortly after birth. In the adult brain, this agranular type of neocortex is found only in the areas 4 ("primary" motor cortex) and 6 ("premotor" cortex) of Brodmann (1909). However, a hardly visible layer IV can still be seen in the prosimian *Indri indri* motor cortex (figure 3.4A). Area 4 is further characterized by the presence of giant pyramidal cells in layer V (figure 3.4A–D), which have not only basal and apical dendrites but numerous dendrites originating from all sides of the cell body (Betz cells). This feature can be seen in all primate species.

In contrast, the primary sensory areas display a six-layered structure with a remarkable development of small round ("granular") neurons in layer IV invading the more superficial cortical layers II–III. This invasion blurs the borders of the superficial cortical layers to such a degree that these layers are difficult to separate. This adaptation of the primary sensory neocortical areas to their massive thalamo-cortical input from the ventroposterior thalamic nucleus (to the primary somatosensory area 3b), medial geniculate body (to the primary auditory cortical area 41 after Brodmann, 1909), and lateral geniculate body (to the primary visual cortical area 17 after Brodmann, 1909) led to the classification of these areas as koniocortical regions. The extreme "granularization" in the superficial cortical layers of the primary sensory areas (figures 3.4E–F and 3.5A–E) differs clearly from the normal density of granular cells in the adjoining secondary and tertiary sensory areas.

The visual system is the dominant functional system in many primates including *Homo sapiens*. This functional specialization is associated with further differentiations of layer IV into the three sublayers IVA, IVB (with the Gennari stripe), and IVC in New World (figure 3.5B), and Old World monkeys (figure 3.5D) as well as *Homo sapiens* (figure 3.5E). The highest level of laminar differentiation within primates is visible, however, in area 17 of the prosimian *Tarsius*. Here, the highest number of sublayers of the inner granular layer is found even when compared with the human brain. This

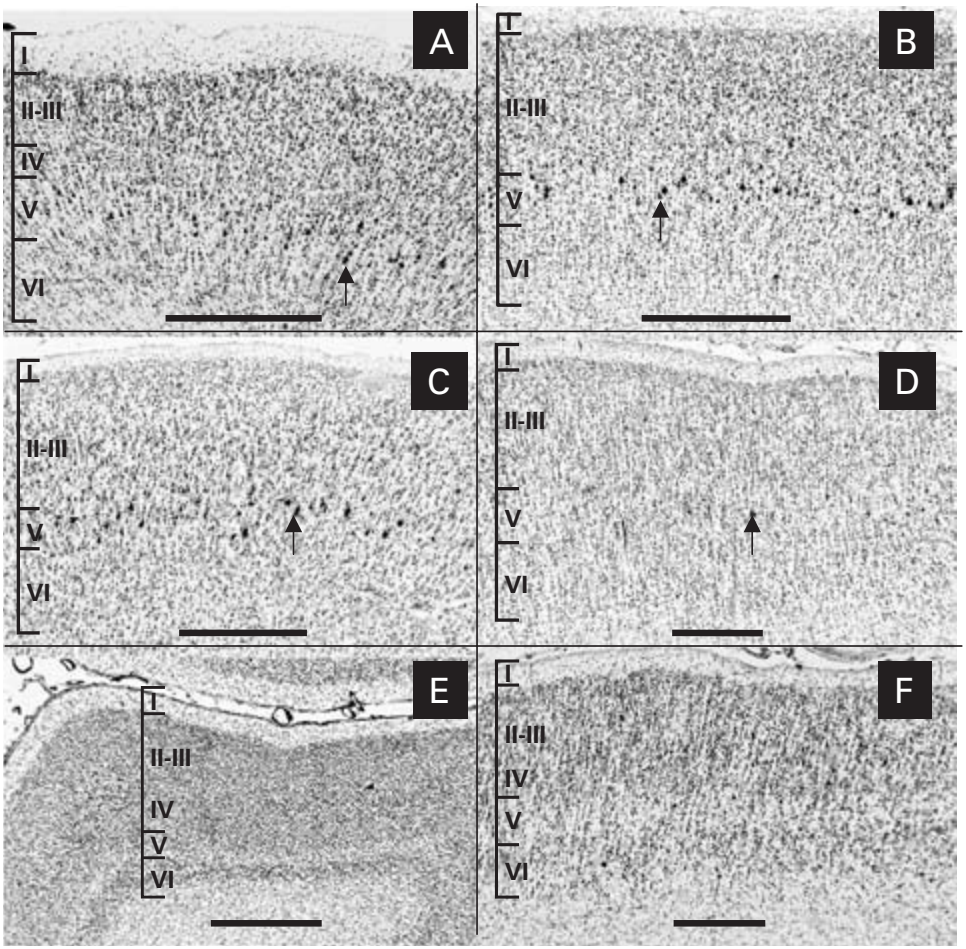


Figure 3.4
Micrographs of the primary motor area 4 (*A–D*) and primary somatosensory area 3b (*E–F*) in the prosimian *Indri indri* (*A*), New World monkey *Alouatta seniculus* (*B*), Old World monkeys *Gorilla gorilla* (*C*) and *Macaca mulatta* (*E*), and *Homo sapiens* (*D*, *F*). Arrows indicate Betz cells in layer V of the primary motor cortex. Roman numerals, cortical layers. Bars: 1 mm.

species has an area 17 which covers more than one third of the whole forebrain, and an eye which is as big as the whole brain. *Tarsius* reflects in the cytoarchitecture of its primary visual cortex, area 17, the excellent performance of its cortical control of visual functions. In contrast, the prosimian *Lepilemur ruficaudatus* shows no subdivision of layer IV (figure 3.5A).

In conclusion, the cytoarchitecture of the cerebral neocortex has a common six-layered structure in all primates and displays a considerable adaption to the differing functional connectivity of sensory and motor systems within each species. Moreover, specific functional adaptations to ecological niches are reflected by the variable cytoarchitecture, e.g., of the primary visual cortex between the different species. Thus, a single species may exhibit a pronounced specialization of cortical structure and function particularly in primary sensory areas, which exceed the degree of specialization found in other species including *Homo sapiens*. However, the flexibility of neural functions and the overall performance of a brain depend more on the degree of intracortical connectivity, and thus, on the evolution of multimodal brain regions than on the selective evolution of one specific sensory system.

Similarities and Dissimilarities Between Mammalian and Human Hippocampus at the Level of Transmitter Receptors

The hippocampus seems to be an evolutionary conservative region in mammalian brains. Its principal regions, Ammon's horn (Cornu ammonis, CA) with its subdivisions CA1–CA4, and dentate gyrus, show the same cytoarchitecture and principal cell types in all species. In contrast to the volume of the neocortex, the volume of the hippocampus shows only minor changes, when analyzed in an allometric study of primates (Stephan, 1975). Is the hippocampus of macaque monkeys, therefore, a model of the human hippocampus?

Studies of the transmitter receptor distributions in the human hippocampus and that of nonhuman primates reveal considerable differences although their anatomical and cytoarchitectonic organizations are well comparable. This becomes particularly evident, when we analyze the receptor distribution patterns in the major intrahippocampal pathway starting with the perforant pathway from the entorhinal cortex to the molecular layer of the dentate gyrus, and leading from there via granular cells of the dentate gyrus and their axons (mossy fibers) to the stratum lucidum of the CA3 region. From here the pathway continues via the Schaffer collaterals of the CA3 pyramidal cells to the CA1 pyramidal cells. The neurotransmission in this system relies on glutamate release and glutamate receptors, but all the other transmitter systems also contribute to the function of the neurons in this system.

The termination field of the mossy fibers in CA 3 and in the hilus of the dentate gyrus show a very high density of kainate receptors both in the human (figure 3.6E)

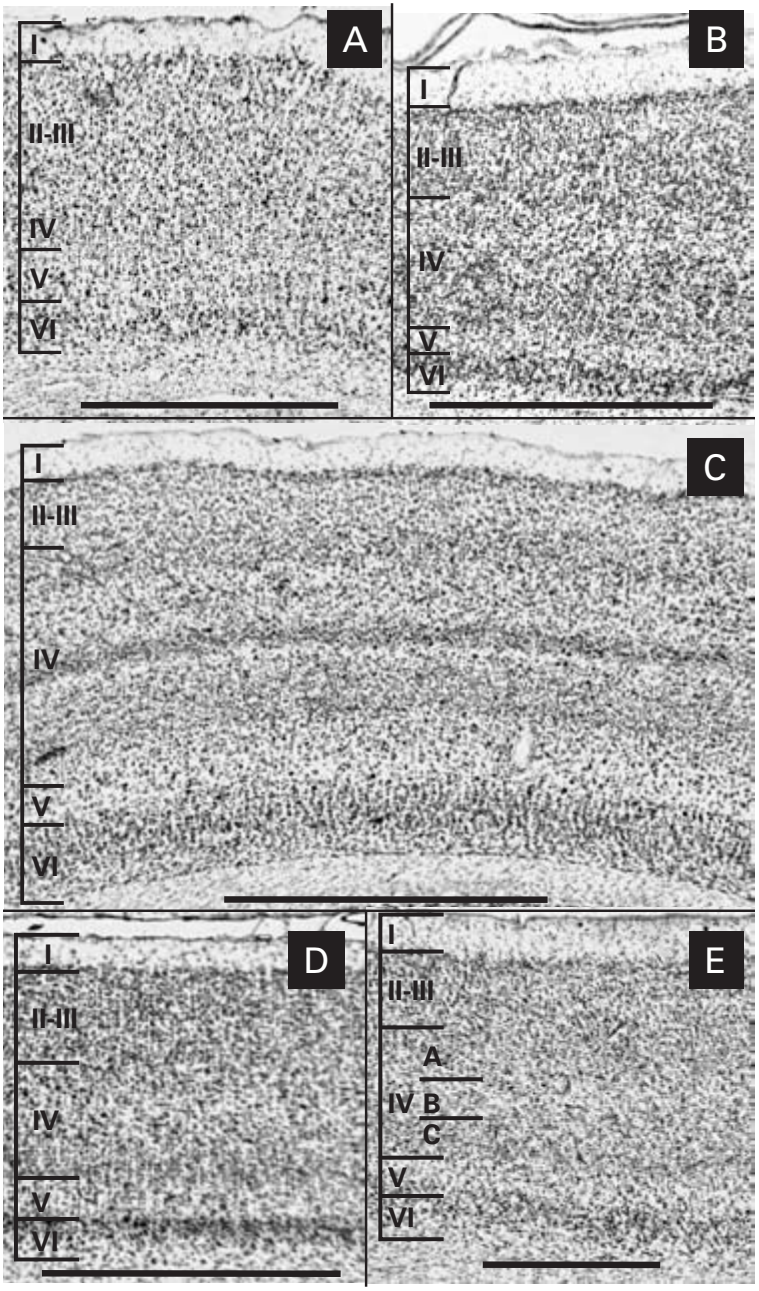


Figure 3.5
Micrographs of the primary visual area 17 (*A–E*) in the prosimians *Lepilemur ruficaudatus* (*A*) and *Tarsius bancanus* (*C*), New World monkey *Alouatta seniculus* (*B*), Old World monkey *Macaca mulatta* (*D*), and *Homo sapiens* (*E*). Roman numerals, cortical layers. Bars: 1 mm.

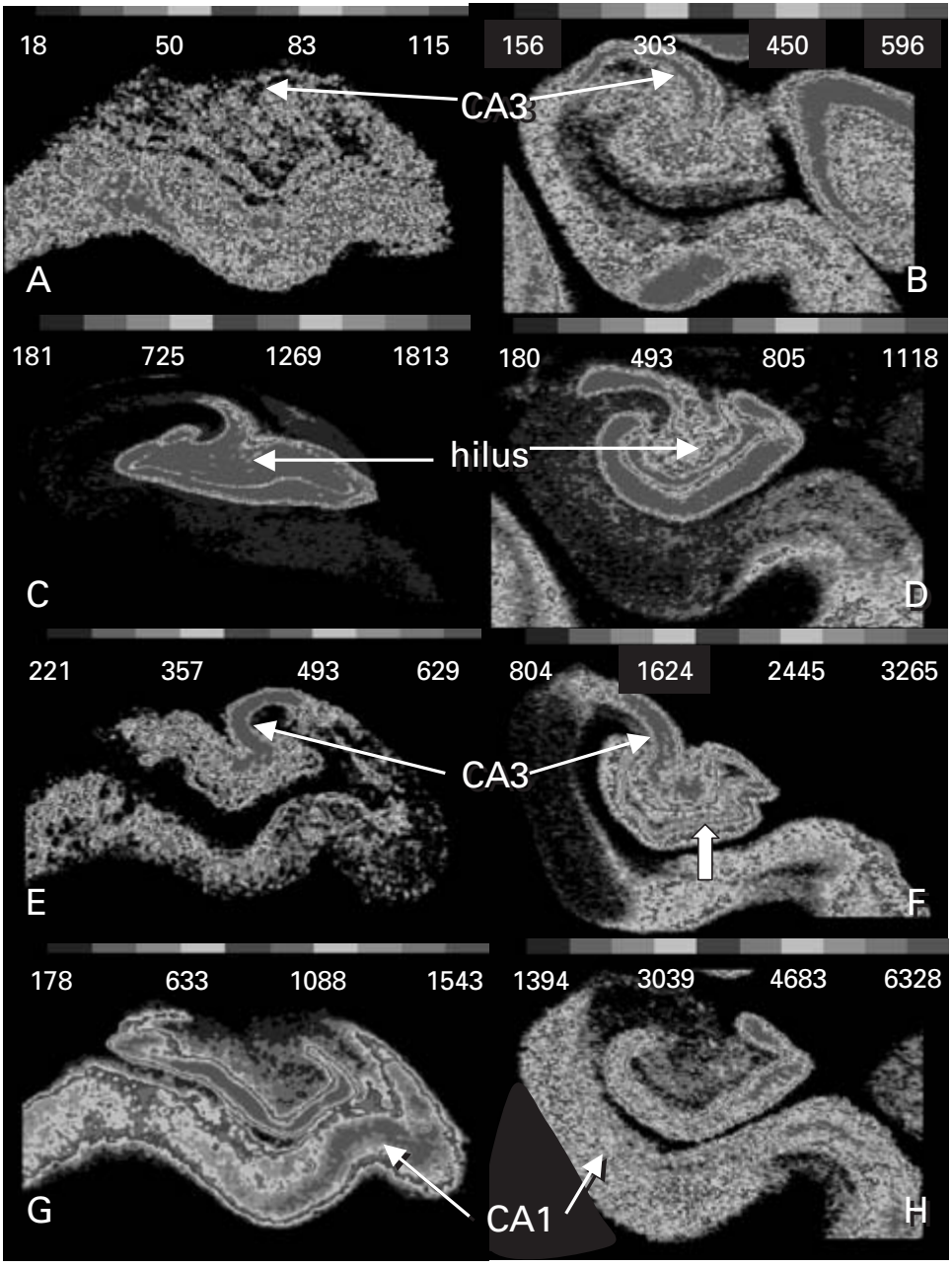


Figure 3.6

Transmitter receptor distributions in the hippocampus of *Homo sapiens* (*A*, *C*, *E*, *G*) and macaque monkeys (*B*, *D*, *F*, *H*). Receptor autoradiography was performed according to standard procedures (Zilles et al., 2002a). (*A* and *B*) Muscarinic M2 receptor; (*C* and *D*) α_1 adrenoceptor; (*E* and *F*) kainate receptor; (*G* and *H*) $GABA_A$ receptor. Receptor densities in fmol/mg protein are shown by color coding. Large arrow indicates high density of kainate receptors in the deeper part of the molecular layer of the dentate gyrus. See plate 12 for color version.

and the macaque (figure 3.6F) hippocampus (plate 12). However, an additional small layer with a high density of kainate receptors is visible in the deepest part of the molecular layer of the dentate gyrus of the macaque (figure 3.6F, large arrow), but not in the human hippocampus. The muscarinic M2 receptor is expressed at a very high density in the CA 3 region of the macaque (figure 3.6B), but not in the human hippocampus (figure 3.6A). The α_1 adrenoceptor shows also differences in its intrahippocampal distribution between humans (figure 3.6C) and macaques (figure 3.6D) particularly in the hilus of the dentate gyrus. Finally, the inhibitory $GABA_A$ receptors are located with a high density in the human CA 1 region (figure 3.6G), but not in the same region of the macaque (figure 3.6H). These results demonstrate considerable changes of the regional and laminar distribution of important signaling molecules in an otherwise evolutionary conservative brain region.

Similarities and Dissimilarities Between Macaque and Human Neocortex at the Level of Transmitter Receptors

Like the above-described cytoarchitectonical observations, the receptor data demonstrate also communalities and differences between the organization of the monkey and human neocortex.

We have studied glutamatergic AMPA, NMDA, and kainate receptors; GABAergic $GABA_A$ receptors; $GABA_B$ receptors and benzodiazepine binding sites; cholinergic muscarinic (M1 and M2) and nicotinic receptors; serotoninergic 5-HT_{1A} and 5-HT_2 receptors; and α_1 and α_2 adrenoceptors, as well as dopaminergic D1 receptors in cytoarchitectonically defined areas of both human and macaque cerebral cortex using quantitative *in vitro* receptor autoradiography (Zilles, Schleicher, Palermo-Gallagher, & Amunts, 2002; Zilles et al., 2002). The autoradiographs reveal the binding sites of the tritiated ligands to their specific receptors. The densities of the different receptor binding sites were measured (fmol/mg protein) and color-coded in autoradiographs of immediately adjacent, 20 µm thick serial sections through deep-frozen whole human and macaque hemispheres (figure 3.7, plate 13). This procedure enables the quantitative analysis of numerous different receptor types in the same cytoarchitectonically defined cortical area.

The primary somatosensory (area 3b), auditory (area 41), and visual (area 17) areas are clearly separated from other neocortical areas by their very high expression of the cholinergic muscarinic M2 receptor in both human and macaque brains (figure 3.7). These primary sensory areas are characterized not only by the highest M2 receptor density (averaged over all cortical layers), but also (not shown here) by very high nicotinic, 5-HT_2, $GABA_A$, and α_2 receptor densities when compared with other neocortical areas. Contrastingly, the glutamatergic AMPA and kainate receptors (not shown here) occur at very low densities in these areas of both species. The adjoining secondary

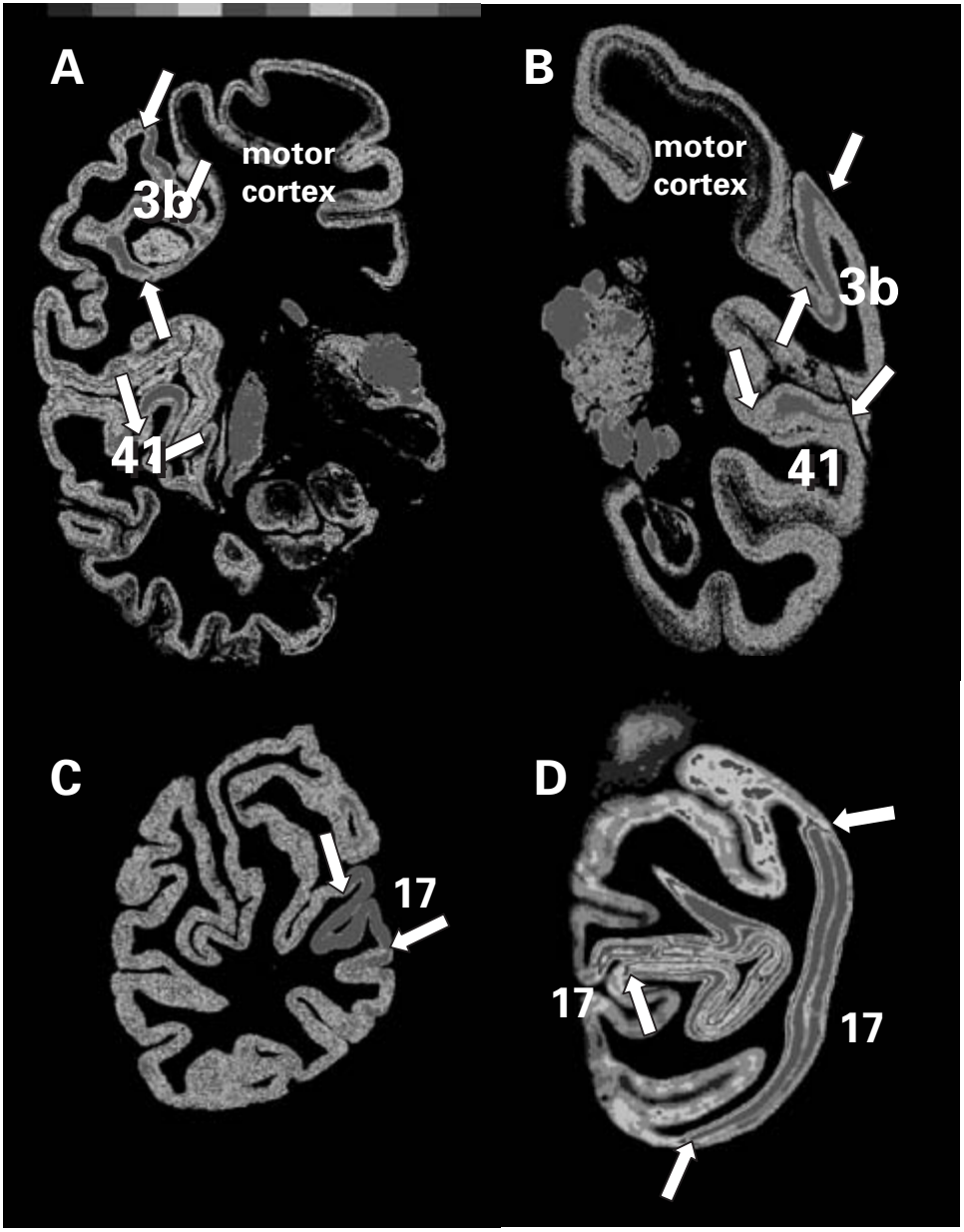


Figure 3.7

Transmitter receptor distributions in coronal sections through a whole human (*A*, central level; *C*, occipital level) and macaque (*B*, central level; *D*, occipital level) hemisphere. Receptor autoradiography was performed according to standard procedures (Zilles et al., 2002a). The distribution of muscarinic M2 receptor is shown. A remarkably high receptor density is found in the primary somatosensory (area 3b), primary auditory (area 41) and primary visual (area 17) cortex. Contrastingly, the receptor density in the motor cortex is very low. Receptor densities in fmol/mg protein are shown by color coding. See plate 13 for color version.

sensory areas differ from their respective primary areas by intermediate to low densities of M2 receptors. The motor cortex and the inferior temporal association cortex are characterized by very low M2 receptor densities (figure 3.7A, B).

When the densities of all receptor in each cortical area are taken into account, multidimensional data sets representing quantitatively the normal balance between different receptors can be displayed as "receptor fingerprints" of cortical areas (Zilles et al., 2002). It is hypothesized that a hierarchical cluster analysis of the fingerprints of various cortical areas reveals organizational principles of the cerebral neocortex at the level of complex systems, e.g., motor, primary sensory, or multimodal association regions (figure 3.8).

A fundamental separation between different receptor fingerprints is found between the visual and all other cortical areas studied here. A second major separation occurs between the prefrontal association areas 10 and 11 (Brodmann, 1909) and all other areas. The third separation takes apart the parietal areas 5 and 7 from the primary motor area 4. The latter area is separated also from the premotor area 6 and the transition area 3a (functionally an area representing proprioception) between motor and somatosensory cortex. The remaining areas, 1, 2, 3b, 41 and 42 are all cortical areas receiving signals elicited by mechanical stimulation of the cochlea (41 and 42) or the skin and muscles (1, 2 and 3b). Thus, the cluster analysis of the receptor fingerprints reveals a functionally meaningful classification of various cortical areas in the macaque cortex, indicating a close association of receptorarchitecture and cytoarchitecture with functional aspects of the neocortex.

The cluster analysis of the receptor fingerprints in the human brain shows similarities, but also remarkable differences when compared with the macaque monkey. As in the monkey cortex the prefrontal areas 10 and 11 cluster together as well as the mechanoreceptive areas 1 and 42, and the visual areas 18v and 18d. In contrast to the monkey condition, the specialization of the primary sensory areas of the human brain seems to reach such a similarity in its receptor distribution pattern that the primary visual (area 17), somatosensory (area 3b), and auditory (area 41) areas form a cluster. Like in the monkey, however, the mechanoreceptive areas 3b (somatosensory) and 41 (auditory) are more closely positioned to each other than to the visual area 17 in the hierarchical tree. A further similarity between monkey and human receptor patterns is the separation of the highly specialized primary motor area 4 from the premotor area 6.

The clearest difference between the receptor fingerprints of macaques and humans is represented by the human cluster of the posterior parietal areas 2, 5, 39, and 40. Human areas 39 and 40 are found in the inferior parietal lobule, where Brodmann's (1909) area 7 is located in the monkey brain. The superior parietal lobule of the human brain comprises areas 5 and 7, where only area 5 is found in the macaque brain (Brodmann, 1909). Area 7 clusters with visual areas 18d and 18v in the human brain,

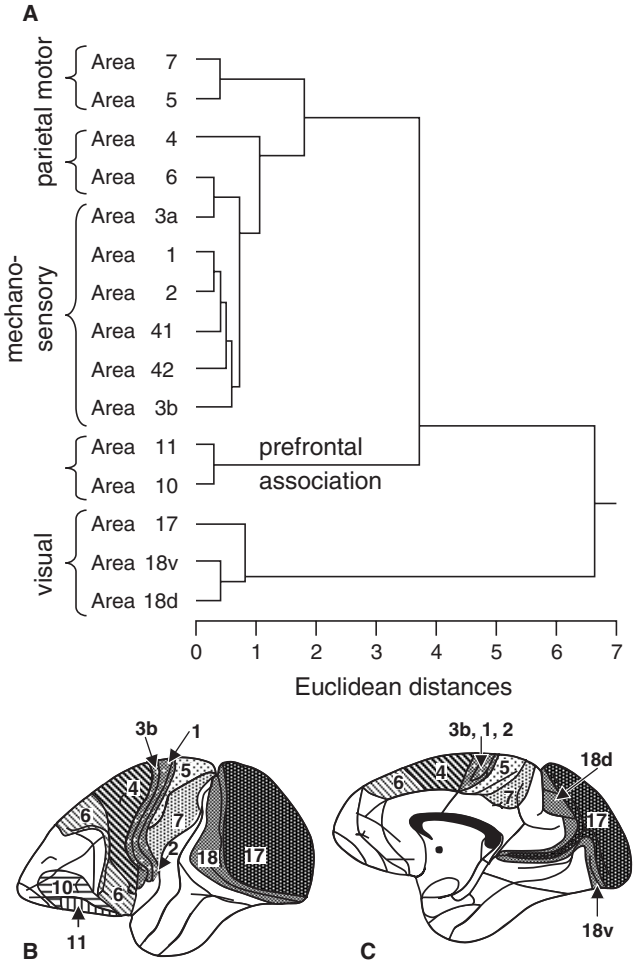


Figure 3.8

Hierarchical cluster analysis (Ward algorithm) of the receptor fingerprints (densities of various receptors averaged over all cortical layers in a cytoarchitectonically defined area) in each of the macaque cortical areas analyzed here (*A*). (*B*) and (*C*) are modified versions of the lateral (*B*) and medial (*C*) views of Brodmann's (1909) map of the monkey cortex showing the localization of each of the analyzed areas. The receptor densities of the following 14 receptors are included: glutamatergic AMPA, NMDA and kainate receptors, GABAergic $GABA_A$ receptors, $GABA_B$ receptors and benzodiazepine binding sites, cholinergic muscarinic (M1 and M2) and nicotinic receptors, serotoninergic 5-HT_{1A} and 5-HT_2 receptors, α_1 and α_2 adrenoceptors as well as dopaminergic D1 receptors.

reflecting the close functional relation of area 7 to a common system in visually guided motor actions. Thus, a considerable reorganization of the posterior parietal region of humans is revealed by the receptor fingerprints when compared with this region in the monkey cortex. A straightforward interpretation of this change in the expression and balance of signaling molecules is presently not possible. However, on the basis of the known function and dysfunction (apraxia, neglect) of the human posterior parietal cortex (Assmuss et al., 2003; Lux et al., 2003; Stephan et al., 2003) we may speculate, that such a change of transmitter receptors in this cortical region may indicate a functional adaption of the expression and balance of signaling molecules subserving extensive tool use and improved egocentric and allocentric spatial attention in humans.

Acknowledgments

This Human Brain Project/Neuroinformatics research is funded by the National Institute of Biomedical Imaging and Bioengineering, the National Institute of Neurological Disorders and Stroke, and the National Institute of Mental Health. Parts of the work were supported by the VW-Stiftung, Brain Imaging Center West, and the DFG.

References

Armstrong, E., Curtis, M., Buxhoeveden, D. P., Fregoe, M., Zilles, K., Casanova, M. F., & McCarthy, W. F. (1991). Cortical gyrification in the rhesus monkey: A test of the mechanical folding hypothesis. *Cerebral Cortex*, *1*, 426–432.

Armstrong, E., Zilles, K., Curtis, M., & Schleicher, A. (1991). Cortical folding, the lunate sulcus and the evolution of the human brain. *Journal of Human Evolution*, *20*, 341–348.

Armstrong, E., Zilles, K., Omran, H., & Schleicher, A. (1995). The ontogeny of human gyrification. *Cerebral Cortex*, *5*, 56–63.

Armstrong, E., Zilles, K., & Schleicher, A. (1993). Cortical folding and the evolution of the human brain. *Journal of Human Evolution*, *25*, 387–392.

Assmus, A., Marshall, J. C., Ritzl, A., Noth, J., Zilles, K., & Fink, G. R. (2003). Left inferior parietal cortex integrates time and space during collision judgments. *NeuroImage*, *20*, 82–88.

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Großhirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Leipzig: J. A. Barth.

Crivello, F., Schormann, T., Tzourio-Mazoyer, N., Roland, P. E., Zilles, K., & Mazoyer, B. M. (2002). Comparison of spatial normalization procedures and their impact on functional maps. *Human Brain Mapping*, *16*, 228–250.

Lux, S., Marshall, J. C., Ritzl, A., Zilles, K., & Fink, G. R. (2003). Neural mechanisms associated with attention to temporal synchrony versus spatial orientation: An fMRI-study. *NeuroImage*, *20*, 58–65.

Mohlberg, H., Weiss, P., Fink, G., Zilles, K., & Amunts, K. (2002). Integration of a new non-linear elastic warping method and SPM99: Verbal fluency and cytoarchitecture. 8th International Conference on Functional Mapping of the Human Brain, 2–6 June 2002.

Schormann, T., & Zilles, K. (1998). Three-dimensional linear and nonlinear transformations: An integration of light microscopical and MRI data. *Human Brain Mapping*, *6*, 339–347.

Stephan, H. (1975). Allocortex. In W. Bargmann (Ed.), *Handbuch der mikroskopischen Anatomie des Menschen*, vol. 4 *Nervensystem*, part 9 (70–77). Berlin: Springer.

Stephan, H., Baron, G., & Frahm, H. (1991). *Comparative Brain Research in Mammals*, vol. 1, *Insectivora*. New York: Springer.

Stephan, K. E., Marshall, J. C., Friston, K. J., Rowe, J. B., Ritzl, A., Zilles, K., & Fink, G. R. (2003). Lateralized cognitive processes entail lateralized task control in the human brain. *Science*, *301*, 384–386.

Zilles, K., Armstrong, E., Moser, K. H., Schleicher A., & Stephan, H. (1989). Gyrification in the cerebral cortex of primates. *Brain Behavior Evolution*, *34*, 143–150.

Zilles, K., Armstrong, E., Schleicher, A., & Kretschmann, H.-J. (1988). The human pattern of gyrification in the cerebral cortex. *Anatomical Embryology*, *179*, 173–179.

Zilles, K., Kawashima, R., Dabringhaus, A., Fukuda, H., & Schormann, T. (2001). Hemispheric shape of European and Japanese brains: 3-D MRI analysis of intersubject variability, ethnical and gender differences. *NeuroImage*, *13*, 262–271.

Zilles, K., Palomero-Gallagher, N., Grefkes, C., Scheperjans, F., Boy, C., Amunts, K., & Schleicher, A. (2002). Architectonics of the human cerebral cortex and transmitter receptor fingerprints: Reconciling functional neuroanatomy and neurochemistry. *European Neuropsychopharmacology*, *12*, 587–599.

Zilles, K., & Rehkämper, G. (1988). The brain with special reference to its telencephalon. In J. Schwartz (Ed.), *Orang-utan Biology* (157–176). New York: Oxford University Press.

Zilles, K., Schleicher, A., Palomero-Gallagher, N., & Amunts, K. (2002). Quantitative analysis of cyto- and receptorarchitecture of the human brain. In A. W. Toga & J. C. Mazziotta (Eds.), Brain Mapping: The Methods, 2nd ed. (pp. 573–602). San Diego: Academic Press.

4 Evolution of the Human Brain and Comparative Paleoanthropology

Jean-Jacques Hublin

In the second half of the nineteenth century, the discovery and identification of fossil hominids resulted in the development of paleoneurological studies dedicated to the first stages of human evolution. One of the pioneers of human palaeontology, Eugene Dubois, who discovered *Pithecanthropus erectus* in Java, devoted a large portion of his career to the study of encephalization in Vertebrates and in Hominids (Dubois, 1923). Paleoneurological studies were conducted in parallel to the development of comparative anatomical studies that attempted to enlighten the peculiarities of the human brain compared to those of other primates. Indeed, the human brain appeared to the first comparative anatomists not to be just an homothetic enlargement of monkey or ape brains. Besides the difference in absolute and relative sizes, the apparent development of the frontal lobes in humans was already emphasized by the anatomists, although conflicting reports were published on this issue (Semendeferi et al., 1997). One of the central question in paleanthropological studies was to determine when the human features first developed in the course of human evolution.

For many decades, changes in the absolute and relative size of the endocast and its different parts have been one of the major issues in paleoanthropological studies. However, although many discussions developed about the occurrence of detailed anatomical features, the brain shape was not directly accessible, but only the morphology and morphometry of the braincase. More recently, thanks to the development of new approaches, other aspects of the development of the brain have been studied. A more physiological approach to brain evolution in hominids has developed. Analyses of growth processes in living primates have unveiled some of the features of the ontogenetic development of the brain and the timing of their appearance during the evolution of hominids. Medical imaging techniques have given access to anatomical details of the basicranium of fossil hominids and allowed more accurate comparisons with living samples.

Size and Shape

In the course of human evolution, the brain witnessed a spectacular increase in size that prolonged the evolutionary tendencies of the upper primates. Regarding absolute size, in less than three million years, brain volume increased from less than 400 ml to roughly 1400 ml. Some of this increase is incipient but already observed in prehuman hominids such as the representatives of the genus *Australopithecus* and *Paranthropus* that display slightly bigger brains than the African apes, our closest relatives (McHenry, 1994). However, the major evolutionary changes occurred in the course of the last two million years within the genus *Homo*. The brain nearly doubled in size, from c. 800 ml in the first genuine humans (*Homo ergaster*) to more than 1400 ml in Late Pleistocene *H. sapiens*, without any spectacular change in body size.

One of the main problems faced by the first paleoneurological studies was related to the fact that although brain size could be somehow estimated via the internal volume of the brain case, body mass of fossil individuals known only by fragmentary remains was difficult and often impossible to estimate. The spectacular increase of the fossil record during the last decades and the development of new proxies for the estimation of body mass allowed to provide a clearer picture of the evolution of the body mass and encephalization quotient in the Hominids in a populational perspective. Recent reviews on this issue (Martin, 1983; McHenry, 1994; Aiello & Wheeler, 1995; Ruff, Trinkaus, & Holliday, 1997) demonstrate that relative brain sizes of extant humans clearly contrast with those observed in other haplorhine primates, which conform with a regular allometric equation. However, this pattern developed rather late during hominid evolution, and early hominids do not differ much from apes. The encephalization quotients in *Australopithecus afarensis* and *Australopithecus africanus* bracket the mean quotients observed in the genuses *Pan* and *Hylobates*. Only "Robust australopithecines" (*Paranthropus boisei* and *Paranthropus robustus*) display a significant increase of the quotient.

Relative brain size increased markedly in the diverging and contemporary branch leading from *Homo habilis* to *Homo ergaster* and *Homo erectus*. It still appears that the increase in absolute size of the brain volume observed along this lineage is initially partly the result of a spectacular increase in the body size. It is mainly with the latest representatives of the genus *Homo* such as *Homo neanderthalensis* and *Homo sapiens* that brain size largely developed independently from body format during the last half million years.

The analysis of general brain morphology mostly concentrated on the general shape, degree of asymmetry (petalia) and attempts to identify sulcal and gyral patterns (Holloway, 1996). Although the absolute and relative sizes of australopithecine brains remain close to ape brains, *Australopithecus* and *Paranthropus* display relatively higher brains. However, arguments have surrounded the notion that, in hominids, brain reor-

ganization has preceded brain enlargement. Several authors (see Holloway 1996, 2001) have recognized a more posterior position of the *sulcus lunatus* in *Australopithecus*, especially in the Taung (South Africa) specimen. This position would indicate that although these hominids displayed a brain size within ape limits, the reorganization of their cerebral cortex was already more humanlike, with a relative increase of the posterior parietal "association" cortex and a reduction of primary visual striate cortex (Brodmann's area 17). On the contrary, Falk (1985) has supported the opposite view that *Australopithecus* still displayed an apelike pattern, and that brain expansion antedated the establishment of a humanlike cortex.

Another area of interest for paleoneurologists is the third inferior frontal convolution. This area displays a peculiar development in specimens considered as early representatives of the genus *Homo* (Tobias, 1987; Holloway, 2001). This observation led Tobias (1987) to the claim that language began with *Homo habilis*. Data is missing or controversial documenting the development of the Broca's area in Australopithecines, but MRI studies conducted on the brain of three great ape species *(Pan troglodytes, Pan paniscus,* and *Gorilla gorilla*) by Cantalupo and Hopkins (2001) suggested that Brodmann's area 44 already display an asymmetry in apes with left-hemisphere dominance. According to these authors, the neuroanatomical substrates for left-hemisphere dominance in speech production predated the human/chimp divergence and is not unique to hominid evolution. However, this view has been challenged by Sherwood et al. (2003), who emphasize the difficulty in precisely defining Brodmann's area 44 in living apes from gross morphologic patterns and the large amount of individual variation observed.

Once more, it is to be noted that the robust australopithecines and the lineage leading to *Homo* show some degree of parallelism in the process of brain reorganization with an expended parietal area and a derived cerebellar morphology also developing in the later *Paranthropus boisei*.

Petalia—that is, asymmetrical projections of occipital and frontal cortex, anteriorly, posteriorly, and also laterally—has been related to human handedness by LeMay (1976). Although cortical asymmetry is observed in nonhuman primates and throughout the vertebrates, the phenomenon is more pronounced in hominids and particularly in *Homo sapiens*. Holloway and Lacoste-Lareymondie (1982) showed that although asymmetries are observed in apes, they rarely display the combination of left occipital/right frontal petalias. According to the same authors, only weak asymmetries are observed in australopithecines. Again, it is with ER1740, considered as an early *Homo*, that a very strong petalial pattern appears. An additional argument for the development of right-handedness in early *Homo* has been provided by Toth (1985) based on the analysis of contemporary Oldowan stone industries at Koobi Fora (Kenya). After a long-term experimental research program of manufacturing and using such stone tools, the polarity of the flaking sequence at Koobi Fora

was proved to display the same patterns as in series produced by modern flakers.

Later specimens of *Homo erectus*, *Homo neanderthalensis*, and *Homo sapiens* have hemispheric asymmetries within the range of variation of extant humans. Their endocast does not usually display very clear sulcal and gyral patterns, and most of the discussions have focused on the general shape of the endocasts that can display some degree of platycephaly in *Homo erectus* and to a lesser extend in Neandertals (Grimaud-Hervé, 1997). In these fossils, Broca's area and the prefrontal cortex are said to have expanded, and the lateral sulcus has moved inferiorly and posteriorly. Overall, in Neandertals, the endocast is not much different from that of recent *Homo sapiens*.

Physiological Cost

In many aspects, the emergence of *Homo ergaster* in Africa and its dispersal in some parts of Eurasia marks a crucial step in hominid evolution (figure 4.1). For the first time, a large, fully bipedal and highly mobile hominid was adapted to the exploita-

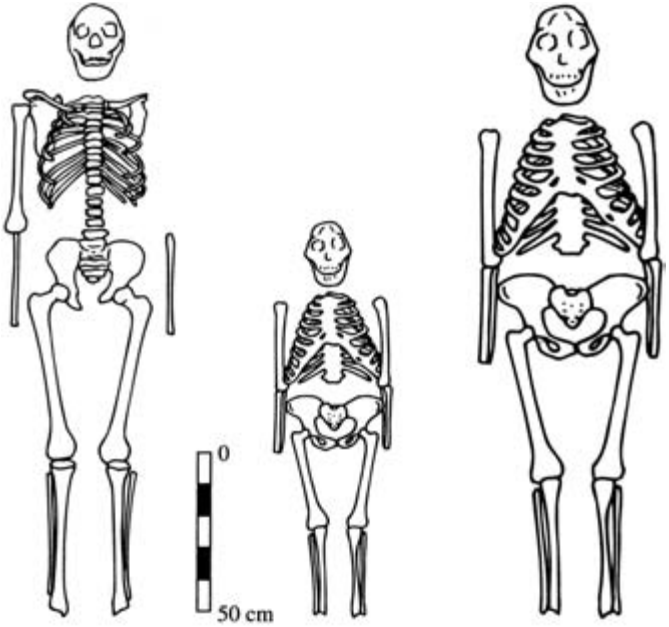


Figure 4.1
(Left) Skeleton of *Homo ergaster* from Nariokotome (West Turkana, Kenya). Middle skeleton of *Australopithecus afarensis* (Hadar, Ethiopia). (R*ight*) skeleton of *Australopithecus afarensis* enlarged to the height of the *Homo ergaster* skeleton. (After Ruff, 1993.)

tion of open landscapes. At this stage hunting and scavenging behaviors became prevalent in humans as well as the large-scale production of stone artifacts.

Evaluations of body mass indicate that the increase in absolute brain volume up to nearly 800 ml 1.8 million years ago (mya) partly resulted from a general increase in body format. However, even the moderate increase in encephalization quotient observed in these hominids (McHenry, 1994) could represent a serious physiological challenge. The brain has been described as an "expensive tissue" (Martin, 1983; Aiello & Wheeler, 1995) from an energetic point of view. Because of its large size and energetic needs, the human brain consumes about 20 percent of the basal metabolism, although it represents only 2 percent of the body mass. This condition strongly contrasts with the one observed in apes and monkeys.

Some of the anatomical and behavioral changes observed in the early representatives of the genus *Homo* are interpreted as resulting from a reorganization of the body energy balance to the benefit of the brain (Aiello & Wheeler, 1995). In this view, the increase in brain size in Hominids was likely made possible by a major shift of a mostly vegetal diet in *Australopithecus* toward highly energetic and concentrated food, namely animal meat and fat, in *Homo ergaster*. This shift is documented by the evolution of dental morphology, by the development of stone artifacts related to the carcasses processing, and by the reduction of the gastrointestinal tract.

The early *Homo* dental morphology displays a reduction of the jugal teeth relatively to the front teeth (Tobias, 1991) in contrast with the large molars of the australopithecines likely needed to process very tough food (Ungar, 1998). Analyzing different microwear patterns, Grine (1981) has suggested that *Australopithecus africanus* had a diet based on fleshy fruits and leaves, while *A. robustus* ate harder, more fibrous foods. However, isotopic analysis have challenged the notion that Australopithecines were mostly fruit and leaf eaters, but rather supported a more broad dietary spectrum in these hominids (Sponheimer & Lee-Thorp, 1999; Lee-Thorp, Thackeray, & van der Merwe, 2001).

The shortening of the gastrointestinal tract in relation to diet change from mostly vegetarian to more carnivorous foods resulted in new body proportions observed in *Homo ergaster* (figure 4.1). In the "expensive tissue hypothesis," this reduction would have allowed to save part of the energy intake that, in Hominids, could have been reallocated to the brain (Aiello & Wheeler, 1995). An alternative hypothesis, called "the maternal investment hypothesis," has been proposed by Martin (1981, 1984). It states that the main limiting factor in brain growth is not so much the adult diet, but the capability of the mother to provide energy to the embryo through the placenta during pregnancy and through breast milk in early life. Energy consumption is indeed higher in neonates than in adults: a newborn's brain consumes up to 60 percent of the energy the baby takes in. More than other primate females, human females are able to provide the embryo with a large amount of energy through the placenta. Actually the

"expensive *tissue hypothesis*" and "maternal investment hypothesis" are not mutually exclusive and could both be partly valid. The human embryo and newborn may use maternal energy sources during the peak period of brain growth thanks to specific adaptations, and a high level of energy consumption by the brain could be later obtained through an appropriate diet and a reduced gastrointestinal tract.

Growth and Development

As mentioned above, developing an embryo and feeding a baby with a big brain represented an evolutionary challenge for hominids from a physiological and energetic point of view. In addition, with the appearance of the first genuine humans, the adaptation to open environments and the improvement of bipedal locomotion observed in *Homo ergaster* also resulted into major changes in body proportions and biomechanics. In contrast with *Australopithecus*, *Homo ergaster* displayed a pelvis of reduced width relatively to the stature and the length of the lower limb that allowed a type of bipedalism different from that of the *Australopithecus*. These longer legs reduced the relative distance between the two *acetabula* and a new morphology of the pelvis resulted in less rotation of the *ilium* during walking. This anatomical structure was not only more efficient for long-distance walking and for running, it created a selective pressure on the width of the inlet and introduced new obstetrical constraints.

This physiological and anatomical dilemma has been solved through adaptative changes of the developmental processes themselves. Humans have a spectacularly enlarged period of growth relatively to other primates and even to apes. This enlarged growth period allows the spread of the costly development of a big brain over a longer period for the mother and for the individual during growth. Furthermore, this general lengthening of the growth is not the only change observed. Humans are born with a brain representing only 25 percent of its adult size (versus 70 percent in a macague or 45 percent in a chimpanzee) and continuing its development at the same embryological rate during the first year of life (figure 4.2). This discrepancy between the development of one organ and the rest of the body is known as "secondary altriciality," a unique pattern distinctive from to the "precocial" reproductive model of some mammals (cetaceans, ungulates, primates, and so on) and the "altricial" model followed by other groups (rodents, carnivores, and so on) (Portmann, 1962). The relatively small brain of human neonates may be a response to the above anatomical and physiological limitations. Extended growth periods and secondary altriciality can be seen as adaptations that allow the birth of an infant at an affordable physiological and anatomical cost by large bipedal females displaying a relatively narrow pelvis. As a result, although in common chimpanzees 85 percent of the adult brain size is acquired one year after birth, in humans a similar proportion is not reached before six years of age.

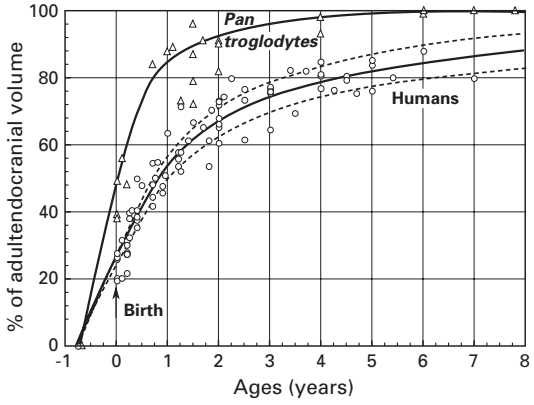


Figure 4.2
Endocranial volume growth in percentage of the adult value in *Pan troglodytes* and extant humans. (After Coqueugniot et al., 2004.)

Second, this enlarged growth period in humans had crucial consequences in term of reproductive pattern and social organization. Human infants are dependent on adults for a much longer time period than in other primates, which induced a peculiar type of mating system and social organization in humans. Their period of learning is proportionally enlarged. However, the most important consequences of the development of secondary altriciality in humans is the fact that, for a long time, brain growth and interactions between specialized cortical areas and the periphery take place while the individual is already interacting with an enriched extramaternal environment and, specifically, with social surroundings. The activity of peripheral somatic areas, such as the orofacial area, interacts with the development of the related sensory-motor cortical areas, resulting in our ability to master articulated language. It is likely that this pattern is also crucial in the development of some mental skills.

When did the enlarged period of growth and secondary altriciality develop in the course of human evolution? Until recently, it was difficult to evaluate the speed of growth of different species of extinct hominids. Recent advances are due to the analysis of tooth microstructures. During dental development, enamel is produced by accretion at a circadian rhythm, which allows precise evaluation of the growth speed of the whole dentition. Analyses conducted on different fossil hominids have indicated that a humanlike enlarged time of growth was not yet established in australopithecines and even in early representatives of the genus *Homo* (Dean et al., 2001). The only well-preserved calvaria of juvenile *Homo erectus* more closely match an ape pattern of brain development rate than a modern human one (Coqueugniot et al., 2004). This suggests that secondary altriciality was established fairly late in the genus *Homo*, perhaps

in the common ancestor of *Homo sapiens* and *Homo neanderthalensis* that both displayed a very large brain and a reduced pelvic inlet size.

Similarities between the adult human skull morphology and that observed in juvenile chimpanzees, as well as the delays observed in the individual development of humans, have led some authors to emphasize the role of heterochrony in the establishment of main hominid features. According to Gould (1977), paedomorphosis, and more precisely neoteny, would have been the main process active during hominid evolution. Neoteny is a deceleration of the development of some or all the organs with a roughly constant time of individual development that leads to the appearance of adult descendants with a juvenile morphology and a size similar to that of the ancestors. In humans, however, the time of development is clearly enlarged, especially for the brain. McNamara and McKinney (McKinney & McNamara, 1991; McNamara, 1997) rejected paedomorphosis as a major process of hominization and favored peramorphosis, and more precisely hypermorphosis, as the main phenomenon. Hypermorphosis prolongs the individual development of the ancestor and results in the increase of size of the adult descendant. These views have been criticized as too simplistic and unable to explain the complexity of the observed processes in human evolution (Gibson, 1991). For Shea (1989), the similarities between adult human skull morphology and those observed in juvenile chimpanzees are only superficial.

Vascularization

The high level of energy consumption and need for improved thermoregulation of the human brain have also resulted in a reorganization of the arterial blood supply and venous blood drainage. Regarding blood supply, preliminary works suggest significant changes with the emergence of the genus *Homo*. In extant humans, the section of the carotid canal displays an allometric relationship with endocranial volume (Braga & Hublin, 1998). This pattern results from the fact that the inner carotid body is highly specialized and provides an almost exclusive blood supply to the nonvisual portions of the cerebral hemispheres that witnessed a spectacular development in humans. This relation does not exist in apes, where a similar development did not occur and where the inner carotid body is also involved in the blood supply of the meninges and the face (Muller, 1977; Diamond, 1991). A scaling analysis of fossil hominids of known endocranial volume (Braga & Hublin, 1998) allowed for determining when the human pattern of the carotidian circulation gained importance. Representatives of *Homo erectus* (in the broad sense), *Homo neanderthalensis*, and early *Homo sapiens* are within the confidence limits associated to the extant human logarithmic linear equation. Specimens assigned to *Australopithecus* and *Paranthropus* display quite different scaling trajectories from both *Homo* representatives and common chimpanzees. It seems, therefore, that the reorganization of the arterial blood supply of the brain did not

occur before 2 mya, and is related with other major adaptive changes observed in hominids during this period.

Another issue that has been much debated in past years is the question of venous drainage and thermoregulation of brain. The increase of brain size in hominid evolution coincides with the adaptation to open, sunny, and hot environments. However, the brain is an organ that requires the highest level of thermal stability, and a rise of more than 4–5 degrees Celsius above normal body temperature can cause convulsions. According to Falk and Conroy (1983), robust and "gracile" australopithecines differed considerably in their cranial blood drainage systems. *Australopithecus afarensis* and *Paranthropus* had an enlarged occipital-marginal sinus that was present in only a small percentage of the *Australopithecus africanus*. Because of the constraints of gravity, bipedalism would have necessitated a rearrangement in cranial blood vessels, producing an enlarged occipital-marginal sinus that delivers blood from the brain into the vertebral venous plexus to be returned to the heart. This would already have been the case for *A. afarensis* and this structure would have remained unchanged in the *Paranthropus* lineage. Falk (1990) developed the view that an alternative system for discharging blood from the skull began to develop in the "gracile" australopithecines (*A. africanus*) and continued in their *Homo* descendants. Along this lineage, the frequency of mastoid emissary veins would have increased in relation with a reduction in the frequency of an enlarged occipital/marginal sinus system and a latter increase in frequencies of parietal emissary veins. Meanwhile, *Paranthropus* would have retained apelike frequencies for numerous features related to blood draining from the cranium, such as a relatively high frequency of multiple hypoglossal canals, relatively low frequency of mastoid foramina, and low frequency of parietal foramina.

Emissary veins are part of a network of veins that cool the brain under conditions of hypothermia in humans (Cabanac, 1995). During exercise, although body temperature can rise to over 40°C, brain temperature can be regulated to 37°C in the cavernous sinus lying on either side of the body of the sphenoid bone. According to the "radiator theory" (Falk 1990), this network developed in relation with ongoing brain expansion from gracile australopithecines to recent humans as, among other factors, brain size was also limited by the necessity to keep its temperature at a level low enough for it to function properly. The "radiator theory" has met several criticisms. Detailed anatomical studies of the basicranial veinous drainage indicate that blood flow can hardly be predicted from studies of bony impressions alone. In Falk's view, *Australopithecus africanus* would have already had the floor plan for a radiator network of veins, while robust australothecines (*Paranthropus*) would have retained a more primitive pattern. However, *Paranthropus* are usually considered to be adapted to a rather open and dry environment. Furthermore, they display higher encephalization quotients than *Australopithecus africanus* (McHenry, 1994; Elton et al., 2001). On a statistical level, Braga and Boesch (1997) did not find any significant difference

between the incidence of divided hypoglossal canals, mastoid canals, parietal and occipital foramina between extant African apes, *A. africanus*, and "robust australopithecines."

Conclusions

The study of brain evolution in fossil hominids meets strong limitations as it is only indirectly accessible. To quote Falk (1986): "In the final analysis, we must remember that despite their usefulness, the information that we glean from endocasts remains (literally) superficial." However, in the last decades, advances in paleoneurology have helped us to unveil some of the processes affecting the evolution of the brain. They partly result from the continuous growth of the fossil record in relation to the constant efforts applied in fieldwork. New specimens are essential to a better understanding of the biology of extinct hominids. In parallel, multidisciplinary approaches need to be developed for the analysis of fossil endocasts and brain evolution. Medical imaging has, in particular, been more and more systematically applied to the exploration of fossil specimens, allowing to access internal anatomical features and to develop a quantitative analysis. Among other things, the recent development of teeth microstructures and the consecutive advances in the knowledge of life history in extant and fossil hominoids have also opened new perspectives on the study of life history of extinct species. Advances in evolutionary studies have also put the emphasis on molecular and genetical aspects aspects of human evolution. Chou et al. (2002) may have identified one important mutation that predated the expansion of human brain. Humans are genetically deficient in the production of the sialic acid N-glycolylneuraminic acid (Neu5Gc), while both N-acetylneura-minic acid (Neu5Ac) and Neu5Gc are found in other mammals. Neu5Gc expression in nonhuman mammals is developmentally regulated and tissue-specific. Interestingly, Neu5Gc is found in most chimpanzee organs, but its expression is selectively down-regulated in the brain. In humans, only or almost only Neu5Ac is found. This is due to inactivation of the gene for the enzyme that converts CMP-Neu5Ac into CMP-Neu5Gc in other mammals. Molecular and paleomolecular studies conducted by Chou et al. (2002) suggest that this gene was inactivated 2.8 mya, before the brain expansion began in humankind's ancestry, 2.1–2.2 mya. Attention and much debate has been also centered around the FoxP2 gene that could be indirectly involved in the production of a complex language and that mutated in the human lineage around 0.2 mya (Enard et al., 2002).

In the early phases of hominid evolution, encephalization appears to have been a limited phenomenon (figure 4.3). In the *Australopithecus* species, brain size remained in the range of observed in apes. In other aspects of their anatomy, however, these hominids seem to separate from apes. Although this has been a matter of debate for many years, there are some clues that the brain already underwent significant changes

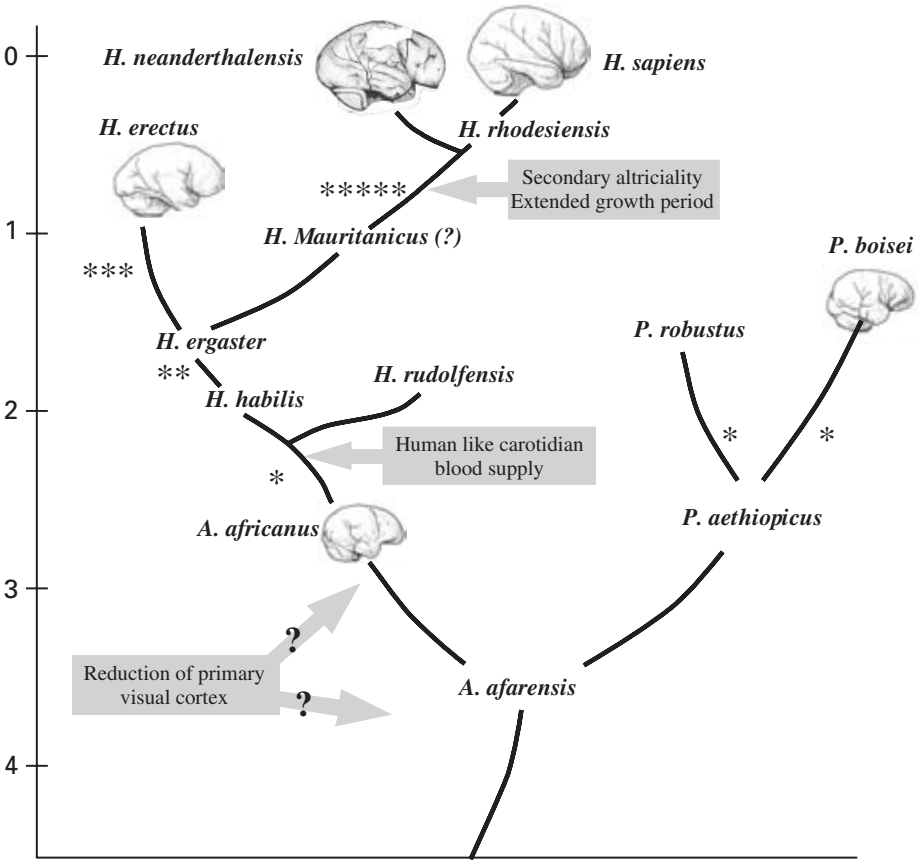


Figure 4.3
Simplified phylogeny of the Hominids with the hypothetical main events in brain evolution.* Increases of the encephalization quotient.

in its organization at this stage of hominids evolution. This is indicated by a possible reorganization of the cortex, with a development of the parietal posterior cortex at the expense of the primary visual cortex, a view that has been supported for many years by Holloway (see Holloway, 1996; Holloway et al., 2001) *contra* Falk (1985). What the inner carotid tells us about brain evolution (Braga & Hublin, 1998) may support the view of the former, as in this respect *Australopithecus* and *Paranthropus* are clearly separate from chimps. Interestingly, some degree of parallelism with a noticeable increase of encephalization is observed in the genus *Paranthropus,* a separate lineage of hominids that is not ancestral to humans and that disappeared 1 mya (Elton et al., 2001).

In contrast, later periods of human evolution are marked by more rapid advances in the evolution of brain size and likely brain complexity. These changes have already been observed in the first representatives of the genus *Homo*, with an increase of the encephalization quotient, a more marked asymmetry (petalia) and the development of Broca's area. However, considering the large amount of homoplasy observed at this stage, the taxonomic and phylogenetic status of the different specimens assigned to *Homo habilis* and *Homo rudolfensis* remains debated (Wood & Collard, 1999a, 1999b) and their endocranial morphology (when known) displays some variability. It is tempting to consider that in relation with the environmental and behavioral changes, about 2 mya, the trend of developing bigger and more complex brains appeared in distinct lineages of hominids (including *Paranthropus*). In many aspects, *Homo ergaster* represents a new biological model of bipedal primate adapted to open environments. It developed toolmaking to a level unrivaled before, and it experienced the first expansion out of Africa. In the species that followed it (*Homo erectus, Homo rhodesiensis, Homo neanderthalensis,* and *Homo sapiens*), one witnesses a clear acceleration of brain evolution. In term of absolute size and encephalization quotient, one observes a very spectacular increase of volume of the endocast. It is quite disappointing to have so little knowledge of changes in complexity and organization of the brain during this period, although in the later stages (*Homo rhodesiensis, Homo neanderthalensis*) the general consensus is that the brain had reached most of its modern human structure.

This evolutionary process might be seen as resulting from an increased pressure of selection on the biological features connected with technical and social skills. But it is also possible that complex interactions between several evolutionary trends, including locomotion and obstetrical adaptations, changes in diet, and body format affected brain development itself. The most recent advances suggest that human life history, with its enlarged period of growth and secondary altriciality, was established rather late in hominids (Dean, 2001). If this view is confirmed by further studies, it will appear that some of the most spectacular changes in the late hominid brain could mainly result from rather simple alterations of growth processes and their control, an evolutionary path that is possible at an economical genetic cost. These changes in the growth pattern also implied concomitant changes in mating patterns and social organization of human groups. Apart from direct effects in the size and organization of the human brain, they also had effects on its maturation process, and likely on the development of some of our mental skills.

References

Aiello, L. C., & Wheeler, P. (1995). The expensive-tissue hypothesis: The brain and digestive system in human and primate evolution. *Current Anthropology, 36*, 199–221.

Braga, J., & Boesch, C. (1997). Further data about venous channels in South African Plio-Pleistocene hominids. *Jurnal of Human Evolution, 33*, 423–447.

Braga, J., & Hublin, J. J. (1998) What do the carotid canals tell us about human brain evolution. *American Journal of Physical Anthropology, suppl. 26*, 112.

Cabanac, M. (1995). *Human Selective Brain Cooling*. Heidelberg: Springer Verlag.

Cantalupo, C. & Hopkins, W. D. (2001). Asymmetric Broca's area in great apes. *Nature, 414*, 505.

Chou, H. H., Hayakawa, T., Diaz, S., Krings, M., Indriati, E., Leakey, M., Paabo, S., Satta, Y., Takahata, N., & Varki, A. (2002). Inactivation of CMP-N-acetylneuraminic acid hydroxylase occurred prior to brain expansion during human evolution. *Proceedings of National Academy of Sciences, 99*, 11736–11741.

Coqueugniot, H., Hublin, J. J., Veillon, F., Houet, F., & Jacob, T. (2004). Early brain growth in *Homo erectus* and implications for cognitive ability. *Nature*, *431*, 299–302.

Dean, C., Leakey, M. G., Reid, D., Schrenk, F., Schwartz, G. T., Stringer, C. & Walker, A. (2001). Growth processes in teeth distinguish modern humans from *Homo erectus* and earlier hominins. *Nature, 414*, 628–631.

Diamond, M. K. (1991). Homologies of the stapedial artery in humans, with a reconstruction ofthe primitive stapedial artery configuration of euprimates. *American Journal of Physical Anthropology, 84*, 433–462.

Dubois, E. (1923). Phylogenetic and ontogenetic increase of the volume of the brain in Vertebrates. *Verhandelingen der Koninklijke Nederlandsche Akademie van Wetenschapen, 35*, 230–255.

Elton, S., Bishop, L. C., & Wood, B. (2001). Comparative context of Plio-Pleistocene hominin brain evolution. *Journal of Human Evolution, 41*, 1–27.

Enard, W., Przeworski, M.,. Fisher, S. E., Lai, C. S. L., Wiebe, V., Kitano, T., Monaco, A. P., & Paabo, S. (2002). Molecular evolution of FOXP2, a gene involved in speech and language. *Nature, 418*, 869–872.

Falk, D. (1985). Hadar Al 162-28 endocast as evidence that brain enlargement preceded cortical reorganisation in hominid evolution. *Nature, 313*, 45–47.

Falk, D. (1990). Brain evolution in *Homo*: The ''radiator'' theory (with peer commentary). *Behavioral and Brain Sciences, 13*, 333–381.

Falk, D., & Conroy, G. (1983). The cranial venous sinus system in *Australopithecus afarensis*. *Nature, 306*, 779–781.

Gibson, K. R. (1991). Myelination and behavioural development: a comparative perspective on questions of neoteny, altriciality and intelligence. In K. R. Gibson and D. Petersen (Eds.), *Brain Maturation and Cognitive Development: Comparative and Cross-cultural Perspectives* (pp. 29–63). New York: Aldine de Gruyter.

Gould, S. J. (1977). *Ontogeny and Phylogeny.* Cambridge, MA: Harvard University Press.

Grimaud-Hervé, D. (1997). *L'évolution de l'encéphale chez Homo erectus et Homo sapiens*. Paris: Cahiers de Paléoanthropologie, CNRS.

Grine, F. E. (1981). Trophic differences between gracile and robust australopithecines. *South African Journal of Science, 77*, 203–230.

Holloway, R. (1996) Evolution of the human brain. In A. Lock and C./R. Peters (Eds.), *Handbook of Human Symbolic Evolution* (pp. 74–116). New York: Oxford University Press.

Holloway, R. L., Broadfield, D. C., & Yuan, M. S. (2001). Revisiting Australopithecine visual striate cortex: newer data from chimpanzee and human brains suggest it could have been reduced during Australopithecine times. In D. Falk and K. Gibson (Eds.), *Evolutionary Anatomy of the Primate Cerebral Cortex* (177–186). Cambridge: Cambridge University Press.

Holloway, R. L. & de Lacoste-Lareymondie, M. C. (1982). Brain endocast asymmetry in pongids and hominids: Some preliminary findings on the paleontology of cerebral dominance. *American Journal of Physical Anthropology, 58*, 101–110.

Lee-Thorp, J. A., Thackeray, J. F., & van der Merwe, N. J. (2001). The hunters and the hunted revisited. *Journal of Human Evolution, 39*, 565–576.

LeMay, M. (1976). Morphological cerebral asymmetries of modern man, fossil man, and non-human primates. *Annals of the New York Academy of Sciences, 280*, 349–366.

Martin, R. D. (1981). Relative brain size and basal metabolism rate in terrestrial vertebrates. *Nature, 293*, 57–60.

Martin, R. D. (1983). *Human Brain Evolution in an Ecological Context: 52nd James Arthur Lecture on the Evolution of the Human Brain*. New York: American Museum of Natural History.

Martin, R. D. (1984). Body size, brain size and feeding strategies. In D. J. Chivers, B. A. Wood, & A. Bilsborough (Eds.), *Food Acquisition and Processing in Primates* (pp. 73–104). New York: Plenum Press.

McHenry, H. M. (1994). Behavioral ecological implications of early hominid body size. *Journal of Human Evolution, 27*, 77–87.

McKinney, M. L., & McNamara, K. J. (1991). *Heterochrony: The Evolution of Ontogeny.* New York: Plenum.

McNamara, K. J. (1997). *Shapes of Time: The Evolution of Growth and Development.* Baltimore: Johns Hopkins University Press.

Muller, F. (1977). The development of the anterior falcate and lacrimal arteries in the human. *Anatomy and Embrology, 150,* 207–227.

Portmann, A. (1962). Cerebralisation und Ontogenese. *Medizinische Grundlagenforschung, 4,* 1–62.

Ruff, C. B. (1993). Climatic adaptation and hominid evolution: The thermoregulatory imperative. *Evolutionary Anthropology, 2*, 53–60.

Ruff, C. B., Trinkaus, E., & Holliday, T. W. (1997). Body mass and encephalization in Pleistocene Homo. *Nature, 387*, 173–176.

Semendeferi, K., Damasio, H., Frank, R., & Van Hoesen, G. W. (1997). The evolution of the frontal lobes: A volumetric analysis based on three-dimensional reconstructions of magnetic resonance scans of human and ape brains. *Journal of Human Evolution, 2*, 375–388.

Shea, B. T. (1989). Heterochrony in human evolution: The case for neoteny reconsidered. *Yearbook of Physical Anthropology, 32*, 69–101.

Sherwood, C. C., Broadfield, D. C., Holloway R. L., Gannon, P. J., & Hof, P. R. (2003). Variability of Broca's area homologue in African great apes: Implications for language evolution. *Anatomical Record, 271A*, 276–285.

Sponheimer, M., & Lee-Thorp, J. A. (1999). Isotopic evidence for the diet of an early hominid, *Australopithecus africanus*. *Science, 283*, 368–370.

Tobias, P. V. (1987). The brain of *Homo habilis*: A new level of organization in cerebral evolution. *Journal of Human Evolution, 16*, 741–762.

Tobias, P. V. (1991). *Olduvai Gorge,* vol. 4, *The Skulls, Endocasts and Teeth of Homo habilis*. Cambridge: Cambridge University Press.

Toth, N. (1985). Archaeological evidence for preferential right-handedness in the lower and middle Pliestocene, and its possible implications. *Journal of Human Evolution, 14*, 607–614.

Ungar, P. (1998). Dental allometry, morphology, and wear as evidence for diet in fossil primates. *Evolutionary Anthropology, 6*, 205–217.

Wood, B., & Collard, M. (1999a). The human genus. *Science, 284*, 69.

Wood, B., & Collard, M (1999b). The changing face of genus *Homo*. *Evolutionary Anthropology, 8*, 195–207.

5 Genes, Brains, and Culture: From Monkey to Human

Jean-Pierre Changeux

It is undeniable that the brain of a monkey or of an ape is not the brain of a human being. Many anatomical and behavioral traits distinguish apes from men including the use of language and access to consciousness. It is equally undeniable, in the genomic era, that such species-specific differences are bound to DNA-encoded mechanisms. Such genetic mechanisms would establish limits, create an "envelope," of the prenatal but also of the postnatal development of the brain and of its cognitive predispositions. We are thus at a dramatic turning point in biological thinking. From the gene sequences stored, *in silico*, in computer memory, all the protein molecules that make up our brain or contribute to its development and plasticity are known or should soon be known. Ultimately, we should be able, in principle, to compute the main features of the organization of our brain from the knowledge of our genetic endowment. At least, we should be able to identify the "signature" of our human nature in terms of DNA sequences and to specify which DNA elements make the difference between the human brain and the monkey brain.

As stated by Karl Zilles (chapter 3, this volume), the human brain is more than simply a large monkey (or ape) brain (figure 5.1). In the course of evolution, the mean brain size increases. In about 2.6–3 million years in human lineage (figure 5.2), the brain volume increases approximately from 457 cm^3 in *Australopithecus africanus* to 1,355 cm^3 in *Homo sapiens*, with intermediate values of 552 cm^3 for *H. habilis* and 1,016 cm^3 for *H. erectus* (Carroll, 2003). In this absolute scale, the macaque brain looks rather tiny with its 63 cm^3, while gorillas and chimpanzees approach the direct human ancestors with respectively 348 and 305 cm^3 (Semendeferi et al., 1997). But if one compares on an allometric scale the different brain compartments by reference to a hypothetical insectivore of the same size, some particular structures of the brain become disproportionately enlarged in the course of brain evolution (see chapter 3, this volume). The paleocortex does not differentially increase from insectivores to humans. On the other hand, the neocortex expands most intensively, largely as a consequence of the exceptional increase of the multimodal association cortex. The number of cortical areas progresses from 10 to 20 in lower mammals (specialized for instance for

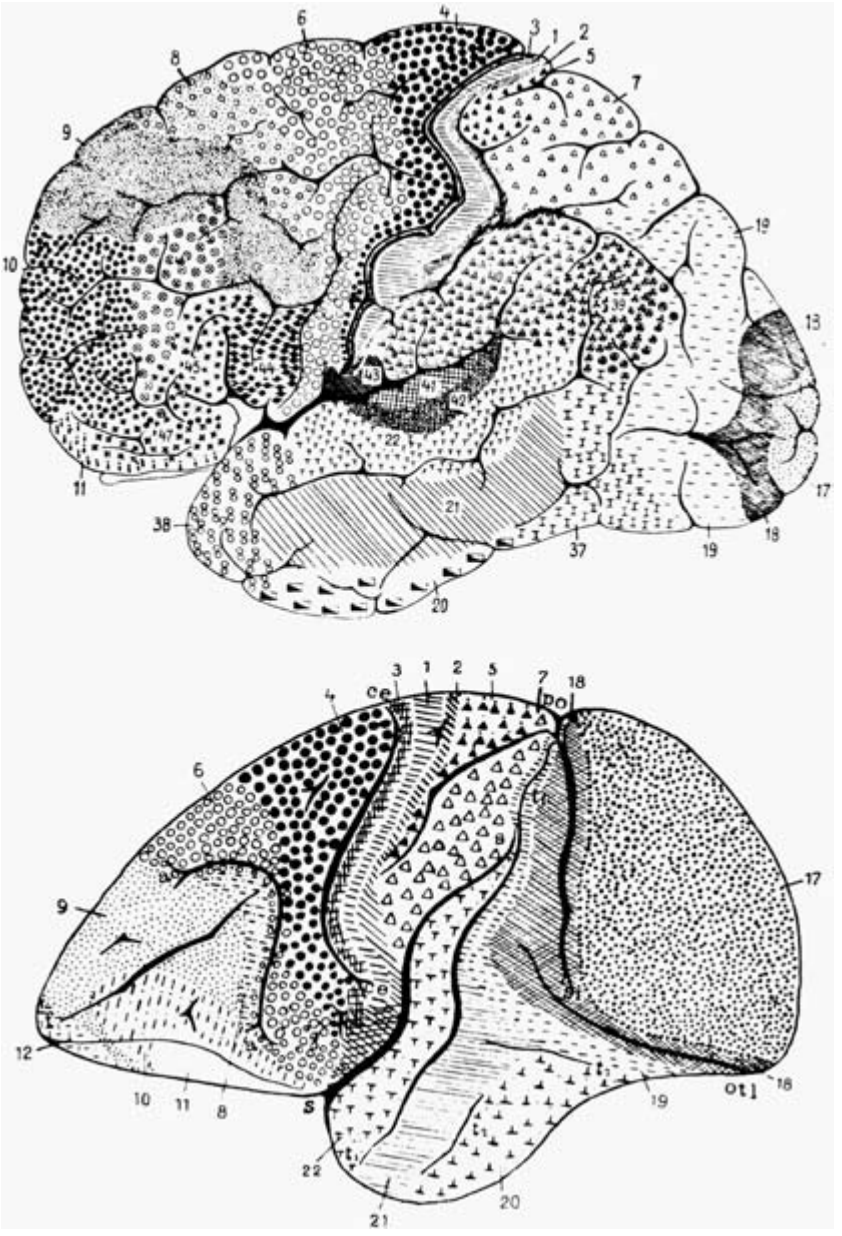


Figure 5.1
Comparison of the maps of the cerebral cortex areas of the human (*top*) and monkey (*bottom*) as published by Brodmann in 1909.

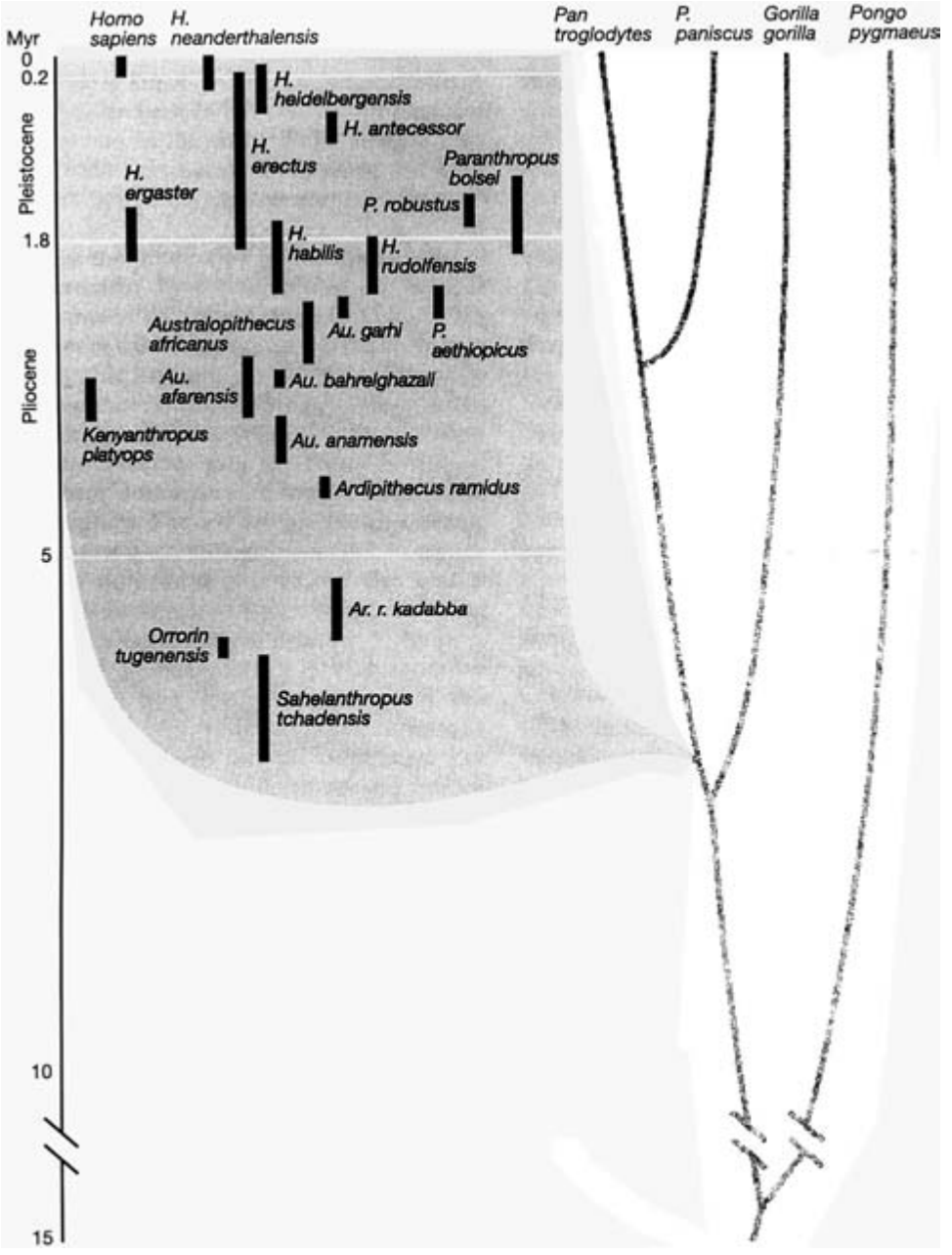


Figure 5.2
Timescale and phylogeny of human lineage (Carroll, 2003).

vision, audition, motor control, and so forth) up to a brain with a very high relative cortical surface in humans, with as many as 100 cortical areas (including association ones) integrated within multiple parallel and hierarchical levels (see Mountcastle, 1998). There is a longstanding notion that the prefrontal cortex, which is mobilized by "higher" cognitive functions, is disproportionately larger in humans (Brodmann, 1912). Recent volumetric analyses by magnetic resonance, however, show a more modest though still significant relative increase of frontal cortex relative volume from 30 percent in the macaque, to 35.4 percent in chimpanzees and 37.7 percent in humans (Semendeferi et al., 2002). One area of the prefrontal cortex, area 10 specifically involved in cognitive functions, is, however, relatively enlarged and more specialized in humans relative to apes (Semendeferi et al., 2001).

Apes and monkeys do not produce and understand speech. Yet, monkeys have been assumed to possess the equivalent of Broca and Wernicke language areas, although without the rich connectivity that characterizes language processing in humans (see Rauschecker, Tian, & Hauser, 1995; Aboitiz & Garcia, 1997; Deacon, 1997). Moreover, a precursor of the human left-right asymmetry of these areas has been found by magnetic resonance imaging (Cantalupo & Hopkins, 2001; Gannon, Holloway, Broadfield, & Brown, 1998). At a more microscopic scale, the cytoarchitecture of the cerebral neocortex has a common six-layered structure in primates, including humans, but varies significantly from one area to the other and from species to species with an unexpected diversity of neurotransmitter receptors patterns (see chapter 3, this volume). Thus an interesting paradox arises. On the one hand, the monkey and human brains share a common plan of organization, yet striking species-specific regional differences unambiguously distinguish them.

Are we in a position to account altogether for these common and different features on the basis of our current knowledge of human genome? This is a challenging issue for the molecular biologist since the number of genetic events which took place in the past million years to achieve such remarkable morphogenetic changes looks astonishingly small!

Genes and Evolution

Before entering the discussion about genes and brains, we should be clear about what we mean by gene. The definition began with the experiments of Mendel on peas and the observations that the inheritance of traits—form, color of the seeds, and so on—follows mathematical rules. This implies the stability of the traits over generations and thus the occurrence of, then, rather enigmatic hereditary "factors." The term *gene* appears in the early 1900s as an abstract concept to account for these hereditary factors, which were, soon after, associated with specific regions of chromosomes by Morgan and his collaborators. Subsequently, Beadle restricted the definition of the

gene to the nature of its product: a protein, thus the famous statement "one gene—one enzyme." Progress in recombinant DNA technologies ultimately lead to a new definition of the gene that is now referred to as the "complete chromosomal segment for making a *functional* product" (Snyder & Gerstein, 2003). The definition includes the expression of a product, which has to be functional, as well as regulatory regions (like promoters) in addition to the protein-coding ones. At the DNA sequence level, structural genes can be identified by start and stop codons that bound a so-called open reading frame within which transcription may take place. Incidentally, the human genome contains a significant contingent of "dead genes" or pseudogenes, which contain frame shift or stop codons in the middle of coding regions, thus preventing them from giving a functional product. Demonstration that a gene is functional requires additional evidence that is not given by straightforward sequence inspection: it is, for instance, the direct demonstration of transcription into mRNA in situ as well as functional inactivation by mutation. Yet, most of the attempts to "annotate" genes in the sequenced genomes do not always satisfy all these criteria simultaneously. It should be emphasized that analyzing the raw nucleotide sequences gives only provisory assignments that often have (or will have) to be revised.

Evaluations of the total number of genes from fully sequenced genomes have nevertheless been published. In man, the current estimate is 30,000–31,000 genes, which are thought to occupy a little more than 1.5 percent of the whole genome (IHGSC, 2001; Venter et al., 2001). Most of the DNA sequence in our chromosomes is thus noncoding. This does not mean, as we shall see, that these noncoding repetitive or intervening sequences have no function. Moreover, proteomics—the study of the proteins actually expressed by a given genome—has to deal with the fact that each individual gene may express as its product more than one protein species. First, it may give several splice variants mRNA and proteins; morever, post-translational modification may create an additional polymorphism (there are about 100 such modifications known). According to O'Donovan et al. (2001) the number of different protein molecules expressed by the human genome is probably closer to a million than to the accepted number of genes, around 30,000. This number is still very small compared to the 50–100 billion neurons in the human brain, and the statements "one gene—one neuron" or even "one gene—one synapse" are blatantly false (table 5.1). There is thus, in sheer number, a striking limitation, a *parsimony*, of available genetic information, not only to code for the body but also, more dramatically, to code for the brain (Changeux, 1983). In order to face this striking parsimony of genetic information, combinatorial mechanisms in gene activities have to be postulated. Today we have access to the full genomic sequences of several major organisms in addition to man (30,000–31,000 genes), for instance: the yeast *Saccharomyces cerevisiae* (6,144 genes) (Goffeau et al., 1996), the plant *Arabidopsis thaliana* (25,706 genes) (Arabidopsis Genome Initiative, 2001), the worm *Caenorhabditis elegans* (18,266 genes) (*C. elegans*

Table 5.1
Nonlinear evolution between the complexity of the genome and the complexity of the brain

Organism	Size of the Genome	Number of Genes	Number of Neurons
Yeast	13.5 Mb	6,144	—
Worm	97 Mb	18,266	302
Fly	165 Mb	13,338	250×10^3
Mouse	3.1 Gb	ca 30,000	40×10^6
Human	3.1 Gb	ca 30,000	$50–100 \times 10^9$

Lander et al., 2001; Venter et al., 2001

Sequencing Consortium, 1998), the fruitfly (*Drosophila melanogaster)* (13,338 genes) (Adams et al., 2000), the fugu fish (Aparicio et al., 2002) and the mouse (27,000–30,500 genes) (mouse Genomic Sequencing Consortium, 2002). This basic information offers the opportunity to compare the full genetic endowment of far distant species and, in principle, to identify the gene sequences that direct the building up of the cells, their assembly into a multicellular organism, and, more specifically, the development of the central nervous system.

A first simple and naïve reasoning is that the evolution of the number of genes parallels the evolution of the complexity of these organisms and, in particular, of their brain. In fact, this is not the case. First of all, the "core proteome," the number of nonredundant proteins coded by the genome of multicellular organisms such as the fly (8,065 proteins) and the worm (9,453 proteins) is only twice that of the yeast single cell (Goffeau et al., 1996; *C. elegans* Sequencing Consortium, 1998; Adams et al., 2000). About 30 percent of the protein coding genes are shared by fly and worm, and nearly 20 percent by fly and yeast. About 1,308 groups of protein coding genes are shared between humans, fly, worm, and yeast, representing respectively 3,129, 1,445, 1,503, and 1,441 protein species (IHGSC, 2001). The shared proteins referred to as *housekeeping* exhibit functions that are common to all eucaryotes: they mediate, in particular, DNA replication and repair, biosynthesis, folding, and degradation of proteins as well as their transport and secretion. These household proteins are essential for cell life, but they do not make the difference between yeast, the fly, and human beings.

Close comparison of the human genome data with those of the invertebrates reveals that only 7 percent of the gene families are unique to vertebrates, and only 12 percent of them are thought to be concerned by the brain. Homologies of many of the genes whose alterations are associated with human neurological diseases, are already found in *Drosophila* (Rubin et al., 2000). The genes that predispose, for instance, to Tay-Sach disease, Duchenne muscular dystrophy, lissencephaly, or fragile X mental retardation possess homologs in the fly genome. Those that cause amyotrophic lateral sclerosis or adenoleukodystrophy are even present in the yeast. Only a few of them look specific

to humans (or vertebrates). They are, for instance, those which cause Charcot-Marie-Tooth or Creutzfeld-Jacob diseases. The fraction of genes that, in the human genome, look proper to humans remains unexpectedly small.

Among the protein coding genes that distinguish vertebrates from invertebrates are the genes engaged in inflammation and immunity, hemostasis, development, cell death, and nervous system function and organization. From the worm and the fly to the human genome, there is thus an increase in the sheer number of members of protein coding gene families that are selectively involved in the development of the nervous system and in neural signalling (IHGSC, 2001; Venter et al., 2001). In particular, as we will see, the genes coding for transcription factors increase dramatically, but so do those involved in the formation of connections, axon guidance (like the ephrins and their receptors), adhesion molecules (such as the proteoglycans), or nerve growth (NGF) or trophic factors (neuregulins). In addition, the genes coding for proteins taking part in the cytoskeleton (like actins and spectrins) expand together with the proteins that compose the myelin sheet. In other words, there is a marked expansion of the proteins involved in building up the neuronal axonal and dendritic arborizations and their interconnections. This is also the case for molecules involved in the propagation of electrical (ion channels) and chemical signals (the opioid peptides are present only in vertebrates). Still, an unexpected nonlinearity (Changeux, 1983) persists in the relationship between the total number of genes and the evolution of brain organization. The worm and the fruitfly genomes are similar in size, with respectively 18,424 and 13,601 different genes. But there are only 302 neurons in the nervous system of the worm and about 250,000 in that of the fly. The number of genes of the flowering plant *Arabidopsis* is even larger (about 25,706), even though it does not have a nervous system. Even more surprising, the number of genes from bony fish to laboratory mouse up to humans is roughly constant, around 30,000.

Moreover, from *Caenorhabditis* to humans, the total number of DNA bases in the genome increases about 30 times from 100 to 2,910 Kb, while the total number of genes increases only three to four times. As mentioned, the coding sequences (referred to as exons) span about only 1 percent of the total genome sequence. Most of the DNA is made up of repetitive or intervening sequences. The size of these noncoding sequences increases with evolution as noted, for instance, between mouse and man genomes (Mouse Genomic Sequencing Consortium, 2002). They may represent traces of the paleontological history but also may serve as active agents in the changes of shape of the genome (see IHGSC, 2001).

In conclusion, if from simpler organims to vertebrates a simple relationship exists between the number of genes and the complexity of the organisms, it is lost in vertebrates, where the estimated total number of coding sequences remains constant while the complexity of brain organization dramatically increases in the course of evolution. How to solve this nonlinearity problem?

Genes and the Morphogenesis of the Brain: From Monkey to Human

The molecular genetics of the early stages of embryonic development in *Drosophila, Xenopus*, chick, mouse, and humans offer one major perspective to deal with both paradoxes of gene parsimony and of nonlinear evolution of gene-brain complexity.

First, in *Drosophila* a variety of genes have been identified that control the Cartesian coordinates of the embryo, the segmentation of the body and the identity of the segments (Driever and Nüsslein-Volhard, 1988). These developmental genes are, for example, the homeotic Hox genes whose mutation in the fly, for instance, causes the transformation of antennae into legs and thus controls the identity of a body segment. Closely related genes are found as well in vertebrate embryos. For instance, there are 160 homeobox domains in humans, 100 in the fly, 82 in the worm, but only 6 in yeast (Venter et al., 2001). Their "nonaccidental resemblance" suggests a common evolutionary origin of the ontogenetic pattern (Arendt and Nübler-Jung, 1999). This is true for the body plan but also for the brain coordinates.

In *Drosophila*, the Cartesian (head-tail, dorso-ventral, right-left) coordinates of the embryo are established very early, before the eggs are laid down. On the other hand, in mammals, they form rather late following the first cleavages of the fertilized egg. In mice, this occurs approximately 6.0 days after fertilization (Beddington & Robertson, 1999). At this stage, the head of the embryonic mouse is labeled by a small set of embryonic cells called the visceral endoderm.

Interestingly these pioneering cells express developmental genes (such as OtX_2) which are also found in *Drosophila* embryo (such as *Orthodenticle*). In addition, the mutation of such genes, for example OtX_2 in the mouse, gives rise to an embryo that fails to develop a normal anteroposterior axis to the extent that the head itself does not form. In other words, OtX_2 and homologs are necessary to make a head both in the fly and in the mouse. Thus despite different timing relationships, the patterns of genes expressed in the course of early development of the body look strikingly similar both in insects and vertebrates (Arendt and Nübler-Jung, 1997, 1999).

The genesis of the vertebrate brain in the course of evolution did not require an entire rebuilding from invertebrates. Gene duplications, possibly two successive duplications of the whole genome, may have sufficed. As was first articulated by Geoffroy Saint-Hilaire with his "unité de plan," a considerable number of structural determinants are saved from our invertebrate ancestors.

An important point to emphasize is that in the course of embryonic and postnatal development, these developmental genes become expressed according to well-defined spatiotemporal patterns, in a hierarchical and parallel manner with cross-regulatory interactions and reutilizations. Such a view of morphogenesis, as a developing network of gene interactions may account, at least in part, for the parsimony paradox. An enor-

mous diversity, indeed, may arise from such combinatorial expression of a limited number of genes.

Also, in vertebrates, a striking asymmetry of the body organs (such as the liver and heart) as well as of the brain hemispheres (in mammals and particularly in humans) does exist. Single genes such as *lefty* in the mouse, or *situs inversus* in humans, together with their protein product (a protein molecule referred to as dynein, known to be involved in flagellar movement) have also been shown to determine the left-right asymmetry of the body plan (Beddington & Robertson, 1999). In all these instances, a few discrete genetic events dramatically change the overall pattern of developmental genes' expression, which directs the three-dimensional plan of the body and thus of the brain. The phenomenon of gene conservation is balanced by the nonlinear expression of a few of them.

In a general manner, at critical stages of embryonic development, "symmetry breakings" take place as manifested, for example, by the development of anteroposterior and dorsoventral polarities, of sharp boundaries between territories, and/or patterns of stripes. Allan Turing proposed in 1952 a theory that accounts for such features. On formal grounds, such defined and reproducible patterns might be generated from a set of chemical substances, or morphogens, which cross-react and diffuse throughout the organism. For instance, gradients of diffusible morphogens are thought to contribute to the unfolding of developmental gene expression, resulting in anteroposterior polarity (Meinhardt & Gierer, 1974). The main actors (but not only ones) are the products of the developmental genes: regulatory proteins, the already mentioned *transcription factors* that control gene transcription at the level of the core RNA polymerase II transcription complex (see Mannervik, Nibu, Zhang, & Leving, 1999). They bind to DNA elements (enhancers or silencers) that lock or unlock the transcription of adjacent structural genes and are themselves often conserved across species. Interplay between morphogens and transcription factors (coactivators and/or corepressors) build up an *intracellular network of gene regulation*, together with membrance receptors and the relevant second messengers. Such sets of molecules may contribute to the "reading" of a gradient of morphogen by an autocatalytic nonlinear switch in both a non-cellularized (Kerszberg & Changeux, 1994) or a cellularized embryo (Kerszberg, 1996). Such reading may even require a rather particular kind of *molecular interconnections* at the level of the transcription factors: for instance, the assembly of molecular partners into allosteric *hetero-oligomers* between one morphogen molecule from the gradient and a transcriptional coregulator now coded by a gene expressed in the embryonic nuclei. *Nonlinear* relationships between transcription factor concentration and morphogenesis may thus emerge from these combinatorial relationships, which include *autocatalytic switches*.

Differential combinations of transcription factors at the level of the promoters would give rise to sharp boundaries of gene expression at definite positions on the

morphogen gradient (Kerszberg & Changeux, 1994). Computer simulations reveal that in addition to the formation of sharp and stable boundaries, the model accounts for the formation of patterns of bands distributed at well-defined positions along the gradient (see also Smolen, Baxter, & Byrne, 2000). The suggested combinatorial and nonlinear networks of transcription factors (Kerszberg & Changeux, 1994) have, since then, been experimentally documented in several examples of *Drosophila* development (Mannervik et al., 1999; Carmena et al., 1998; Halfon et al., 2000) among others (Davidson et al., 2002). Among other predictions, the model accounts for the displacement of the band pattern experimentally observed in *Drosophila* embryo as a function of transcription factor levels (Hoch, Seifert, & Jäckle, 1991). It illustrates how the activation of a given gene may depend on the "context" of genes expressed through the network of transcription factors which link them together.

Along these lines, one may note that what makes the difference between yeast and the nematode as a "minimal" multicellular organism are genes coding for transcription factors (e.g., 270 nuclear hormone receptors), protein-protein interaction domains (e.g., 156 POZ domains), or signal transduction domains (e.g., 11 phosphotyrosine binding domain) (Chervitz et al., 1998). In humans, an even more dramatic expansion of transcription factor genes takes place. For instance, the number of transcription factor domains referred to as CH2H2 zinc fingers increases from 771 in the fly to the huge number of 4,500 in humans. Consistent with our hypothesis of *transcription factor networks* formation, many such factors contain domains that are involved in the oligomerization of transcription factors. Some of them, like KRAB or SCAN, are even absent from the fly or worm genomes. Others result from the reshuffling or "accretion" of already existing invertebrate domains. These gene domains are expected to increase the combinatorial partnering of the transcription factors (Venter et al., 2001). Many of these new genes that I referred to in the past as "communication genes" (Changeux, 1983) are dedicated to the formation of patterns of gene interactions. Should we then say that what make us human lies in transcription factor genes and in the promoters to which they combine?

Many developmental genes are expressed in the nervous system. They are parts of a still largely uncharacterized population of genes concerned with brain morphogenesis from monkey or humans. The concept mentioned for embryonic development may apply as well to brain morphogenesis, in particular to its very early stage of formation referred to as neurulation (see Kerszberg & Changeux, 1998). Interestingly, the process of neurulation differs strikingly in invertebrates and in vertebrates. In the first case, the neuroblasts delaminate from the neural ectoderm to progressively form the solid ganglion chain of the adult (which may reach up to 520 million neurons in *Octopus*). In the second, the neural plate invaginates *en bloc* to form a hollow neural tube, which, as such, may facilitate a dramatic growth of the central nervous system through a surface expansion as is observed all the way from cyclostomes to mammals,

primates, and humans. Such transitions may not require a large number of molecular changes at the gene transcription level. These genetic events may, for instance, affect transcription factor switches, which themselves regulate cell motion (Kerszberg & Changeux, 1998) and cell adhesion (see Edelman, 1988). As a consequence, either the whole neural tissue (neural plate) infolds into a tube (vertebrates), or individual neuroblasts delaminate yielding a solid nervous system (invertebrates). This illustrates in still highly hypothetical terms how a few gene changes may contribute to the critical transition between the invertebrate and vertebrate nervous system. A similar reasoning may apply to the shift of the position of the nervous system from ventral to dorsal which distinguishes invertebrates from vertebrates.

Even less understood is the expansion of the cerebral cortex that took place in the course of vertebrate brain evolution, in particular from monkey to man (see Mountcastle, 1998; Changeux, 1983). The number of neurons per cortical column is rather uniform throughout the vertebrates (Rockell, Hiorns, & Powell, 1980). Thus the surface area of the cortex, i.e., the number of columns, appears as the primary target of the evolutionary changes (Rakic, 1988). The gestation lasts 21 days in the rat, 165 in the macaque, and 280 in humans, and the rapid phase of synaptogenesis (which starts two months before birth in macaque and four to five months after birth in man) lasts 136 days in macaque and 470 days in humans. One may further speculate that the fast expansion of the frontal lobe and parietotemporal areas, which contributed to the evolutionary origins of *Homo sapiens*' brain, resulted from the exceptionally prolonged action of some still unidentified developmental genes (Changeux, 1983), the genomic evolution underlying this process engaging a rather small set of genes or regulatory sequences. An important consequence of the increased surface of the cerebral cortex is the increased possibility of establishing connections between cortical neurons.

As discussed (Changeux, 1983), this translates into an increase in the mean number of connections per neurons, with a consequent burgeoning of the dendritic and axonal trees. For instance, the average number of synapses per neuron increases from 2,000 to 5,600 in the monkey up to 6,800 to 10,000 in humans (Bourgeois, 2003; see Ramon y Cajal, 1910–1911). This is especially true for the prefrontal association cortex, which differentially expands in the human lineage. Indeed, according to Elston (2003), human prefrontal pyramidal cells, which have, on the average, 15,100 spines in their basal arbors, show 72 percent more spines than those in macaque, 3.8 times more than in marmoset, and 7.5 times more than those in *Aotus*. Also, the differential increase in complexity concerns more the prefrontal association cortex than the primary cortical areas: for instance, the prefrontal pyramidal neurons are on the average 16 times more spinous than those in V1 (Elston, 2000). Moreover, von Economo's pioneering observations on the architectonics of cortical areas (1929) underlined that layer II–III pyramidal neurons are more abundant in prefrontal and

parietotemporal association areas than in the rest of the cerebral cortex. Since the layer II–III neurons are known to possess long-range cortico-cortical and callosal axons, the differential expansion of the prefrontal cortex is accompanied by an increase of long range connectivity and in a general manner of the white matter (Olesen, Nagy, Westerberg, & Klingberg, 2003). These convergent findings are thus consonant with the proposal that the long-range cortico-cortical connexions contribute to the neural architectures of a global "conscious workspace" (Dehaene, Kerszberg, & Changeux, 1998; Baars, 1988) in which performances increase from monkey to human (see Dehaene, Sergent, & Changeux, 2003), and which develops in the course of child development (Olesen et al., 2003).

Can we identify the particular genes that cause this neocortical surface expansion? Close comparison of the gene repertoires between mouse and human and among mammals in general reveals a high degree of conservation (Mouse Genomic Sequencing Consortium, 2002). Even more surprisingly, some particular gene families—like the olfactory receptor gene family—show a decrease of functional genes in humans compared to chimpanzees. Humans have more than 1,000 olfactory receptor genes, but only approximately 40 percent have an intact coding region and are therefore functional. In contrast, the fraction of intact olfactory genes in the genomes of the great apes is significantly greater (68–72 percent) (Gilad, Bustamante, Lancet, & Paabo, 2003). We have to consider the unexpected possibility that from early mammals to humans, the number of genes may have decreased!

Overall, looking for genetic changes in the coding sequence of proteins from chimpanzees to man, according to Carroll (2003) the most extensive comparison indicates an average substitution level of about 1.2 percent in single-copy DNA. In addition, gene duplications, chromosomic reorganizations, insertions, and deletions might have to be considered. The total number of amino acid replacements that contributed to protein evolution in the human lineage is estimated to be of the order of 200,000. Yet, one does not know which fraction caused functional changes. The number of so-called adaptative substitutions would then be even smaller: between 10,000 adapatative substitutions in human genes and their regulatory regions (Carroll, 2003). Are all of them unique to humans and their cognitive functions? Which ones are candidates for being causally implicated in past human evolution? In a recent provocative report mutations of the gene FOX P2 (forkhead box P2) were found associated with speech and language disorder (Lai et al., 2001; Enard et al., 2002; Lai et al., 2003). They cause severe articulation difficulties accompanied by linguistic and grammatical impairment. Is this the long-expected "gene for language"? In fact, this is not the case. It is found in other species, and the human FOX P2 differs from the gorilla and chimpanzee sequence at just two residues and from the mouse sequence at four residues (Enard et al., 2002). FOX P2 is a transcription factor, and one may anticipate that, as mentioned above, it contributes to complex regulatory networks of gene interactions.

Its inactivation may, among other factors, cause a disorganization of the network of transcription factors (see Kerszberg & Changeux, 1994, 1998).

Even if the number of genetic changes that took place in the human lineage looks rather small, even if it ultimately boils down to a few thousand, the relationships between these genes and the evolution of neural architecture from monkey to humans cannot be thought of as straightforward. Multiple nested networks of interactions are expected to develop among gene products, in space and time, in the course of human body and brain formation. As a consequence, single-gene actions are expected to be altogether pleiotropic: they affect multiple aspects of the functional organization of the brain (and even of the body). Furthermore, they are context-dependent, and are therefore anticipated to differ in different tissues or organs and at different stages of development. To talk about "language genes" or "genes of intelligence" no longer seems appropriate. But to decipher the regulatory network of transcription factors that lead to the differential expansion of the neocortex and, more specifically, of the prefrontal cortex begins to appear as a plausible challenge for the next decade.

The Activity-Dependent Epigenesis of Neuronal Networks and the Origin of Culture

In the preceding chapters, the paradox of nonlinear evolution between genome and brain complexity has been discussed, exclusively, on the basis of gene interactions. A straightforward prediction of this scheme is that genetically identical, individuals—identical twins or clones—should be neurally identical. In reality, this is not the case. An important variability of the organization of the nervous system exists between genetically identical individuals. For instance, in vivo measurements of the *planum temporale* by magnetic resonance and the results of behavioral tasks collected with monozygotic co-twins discordant for manuality yielded convergent results. The right-handers showed leftward hemispheric asymmetry, whereas the left-handers lacked asymmetry (Steinmetz et al., 1995; Eckert et al., 2002). Early "epigenetic" events take place during embryogenesis and contribute to a significant variability in the development of the anatomo-functional laterality of the cerebral hemispheres. At the neuronal level, a phenotopic variance has been identified in cloned parthenogenetic individuals from the water flea *Daphnia magna* and the fish *Poecilia formosa* (see Levinthal, Macagno, & Levinthal, 1976). At the electron microscope level, the details of the axonal or dendritic branching of an identifiable neuron differ among genetically identical individuals, and the within-individual variability between left and right arborization is smaller than that found from between individuals. In other words, "cloned" individuals are not neurally identical, even in species with a fixed number of identifiable nerve cells. In a mammal such as the mouse (the situation might be even more extreme in humans) the number of cells is much larger, and there are no longer identifiable single cells. Despite common principles of neural architecture

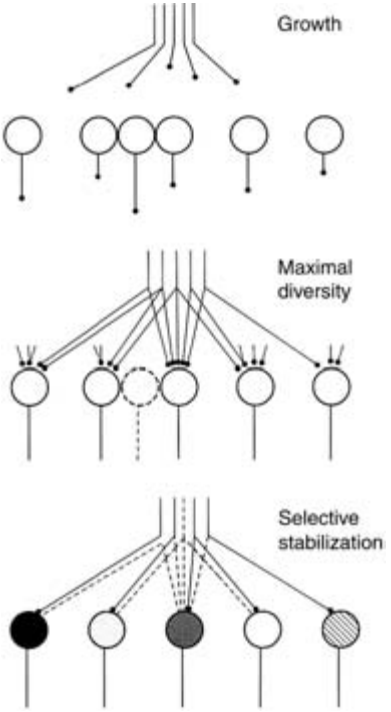


Figure 5.3
The model of epigenesis by selective stabilization of synapses. (From Changeux, 1983.)

delimited by a species-specific genetic envelope, the individual variability of the fine anatomy observed between individuals from genetically homogeneous lineages increases dramatically.

One plausible solution (among many others) is that the state of activity of the developing nervous system contributes to the organization of the adult network by trimming up synapse formation at sensitive periods of development. The model suggested (Changeux, Courrège, & Danchin, 1973; Changeux, 1983; see also Elliott & Shadbolt, 1998) (figure 5.3) posits that during synapse formation the genetic envelope controls—in addition to the division, migration, and differentiation of cell categories—the behavior of the growth cone, the outgrowth and formation of widespread connections, the recognition of the target cells, and the onset of spontaneous activity; it also determines the structure and function of the molecules that enter into the architecture of the synapse, the rules governing their assembly, and their evolution in relation to the activity of the network. Yet, at sensitive periods of circuit development, the phenotypic variability of nerve cell distribution and position, as well as the exu-

berant spreading and the multiple figures of the transiently formed connections originating from the comings and goings of growth cone behavior, introduce a *maximal diversity* that is then reduced by the "active" selective stabilization of some of the labile contacts and the elimination (or retraction) of the others at subsequent stages of development. The crucial hypothesis of the synapse elimination model (Changeux, Courrège, & Danchin, 1973) is that the evolution of the connective state of each synaptic contact is governed globally and within a given time window by the overall signals received from the cell on which it terminated. In other words, the activity of the postsynaptic cell regulates the stability of the synapse in a *retrograde manner.*

The contribution of neural activity (evoked and/or spontaneous) in the formation of cortical circuits has been documented since the classic experiments of Wiesel and Hubel (1963), which demonstrated the important role of visual experience in fixing the organization of ocular dominance columns (see Katz & Shatz, 1996). The exuberant sprouting and proliferation of axonal branches, accompanied by limited though critical elimination of collaterals, has been vizualized at different locations along the visual pathway (retinogeniculate, thalamocortical, and pyramidal cell arbors) (see Sretavan & Shatz, 1986; Katz & Shatz, 1996) at sensitive periods of development. It is clear that the state of activity of the developing cortical circuit controls synaptic evolution by gestation and continues actively during the first two postnatal years. In humans, the *mean synaptic density* reaches a maximum near postnatal age 3 months in the auditory cortex but only *after* age 15 months in the late maturing middle frontal gyrus (Huttenlocher & Dabholkar, 1997). A phase of *net decrease* in mean synaptic density occurs subsequently in childhood, where it ends earlier in auditory cortex (until 12 years) than in prefrontal cortex (until mid-adolescence). Most likely, this global evolution of the mean synaptic density represents the summation of "nested" waves of synapse outgrowth and elimination, the later process becoming dominant and thus visible in global measures of net synaptic density late in development (Huttenlocher & Dabholkar, 1997; Bourgeois 2001, 2003).

The example of the neuromuscular junction is particularly simple since only a single synaptic contact persists in the adult, while in the newborn rat each fiber receives four or five *active* motor axon terminals. As the rat begins to walk, the number of functional terminals progressively decreases until for the adult only one is left. The state of activity of the innervated muscle controls this elimination (e.g., Benoit and Changeux, 1975, 1978; see Sanes and Lichtman, 1999). Similar regressive phenomena have also been documented at the synaptic level in other systems such as the sympathetic ganglia (Purves & Lichtman, 1980) or the climbing fiber Purkinje cell synapse in the cerebelum (see Crépel, Mariani, & Delhaye-Bouchaud, 1976; Changeux & Mikoshiba, 1978; Kano et al., 1997). For the latter, a mutation which inactivates a specific neurotransmitter receptor (the type 1 metabotropic glutamate receptor mGluR1) delays the regression of supernumerary climbing fiber innervation (Kano et al., 1997).

Among the various consequences of our modeling approach, a simple one has been to look for the molecular mechanisms of synaptic outgrowth, stabilization, and elimination and their regulation by neural activity. At the presynaptic level, neurotrophins like NGF, BDNF, NT4 and several others (see Levi-Montalcini, 1987; Barde, 1990) have become plausible candidates for retrograde signals in the activity-dependent-synaptic outgrowth and selection (Thoenen, 1995; Katz & Shatz, 1996). For instance, in vivo intracortical infusion of diverse trophins prevents the shift of ocular dominance in favor of the nondeprived eye (Maffei et al., 1992; Carmignoto et al., 1993) on the formation of ocular dominance columns (Cabelli, Hohn, & Schatz, 1995). Also, rapid and opposite effects of BDNF and NGF have been demonstrated on the whisker barrel representation of the rat somatosensory cortex. Moreover, neurotrophins modulate synaptic strength within minutes in vitro at cultured neuromuscular synapses (Lohof, Ip, & Poo, 1993) and GDNF overexpression in the mouse delays the elimination of supernumerary motor axons and causes hyperinnervation of neuromuscular junctions (Nguyen et al., 1998). Formal models of synapse-selective stabilization based on competition for limited stocks of trophic factors have thus been developed (Gouzé, Lasry, & Changeux, 1983; Kerszberg, Denaene, & Changeux, 1992; Elliott & Shadbolt, 1998), giving even more plausibility to the theory. A molecular genetic approach of activity-dependent synapse selection thus becomes plausible (Schaeffer, Duclert, Huchet-Dymanus, & Changeux, 1998; Schaeffer, de Kerchove d'Exaerde, & Changeux, 2001; de Kerchove D'Exaerde et al., 2002), raising the issue of the relative contribution of these epigenetic processes to the complexity of the brain (see Changeux, 1983).

In mammals a striking relationship exists between the length of gestation and the complexity of the brain: 21 days in the rat, 65 in the cat, 165 in the macaque, and 280 in man. Synaptogenesis in the neocortex of the newborn has been subdivided by Bourgeois (2001) in four phases. Phases 1 and 2 correspond to low density early phases. Phase 3 is rapid and accounts for the overall maximal density of synapses in the cortex. It takes place postnatally in the rat and cat (2–19 days respectively *after* birth) but starts before birth in the monkey and human (2 months and 4–5 months respectively before delivery). But, most significant for us is the increased duration of this phase 3: 14 days in the rat, 30 in the cat, 136 in the monkey, and 470 in the human, thus an increase by at least thirtyfold from rat to human. Bourgeois (2001) has noted that the early phases of synaptogenesis 1 and 2 are independent of the environment, while the subsequent phases 3 and 4 require the interaction with the outside world to develop normally. There is thus a postnatal extensive prolongation of the experience-sensitive phases of synapse formation in the infant brain, which has major consequences. The brain of the newborn human is exposed to the physical, social, and cultural environment for extended period of times compared to the monkey. As a consequence, it becomes accessible to "cultural imprints" such as those involved in spoken

and written language learning. A culture may thus develop. As discussed (Changeux, 1983; Carroll 2003), a few mutations, for example on regulatory genes controlling hormonal release, may have caused this increased duration and thus simply made possible the onset of human culture.

Brain epigenetic capacities to store stable representations of the outside world give human beings the opportunity to create an artificial world of cultural objects at the social level. In other words, the origin of culture and of its transmission from generation to generation lies in the considerable increase of synapse numbers and multiple nested processes of activity-dependent synapse selection that take place postnatally in the human brain. This epigenetic evolution also has another consequence: it permits the diversification of the cultures that human beings have developed throughout their recent history. In other words, the postnatal epigenetic evolution of brain connectivity opens the way to cultural evolution.

References

Aboitiz, F., & Garcia, R. (1997). The anatomy of language revisited. *Biological Research*, *30*, 171–183.

Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., et al. (2000). The genome sequence of Drosophila melanogaster. *Science*, *287*, 2185–2195.

Aparicio, S., Chapman, J., Stupka, E., Putnam, N., Chia, J. M., Dehal, P., et al. (2002). Whole-genome shotgun assembly and analysis of the genome of Fugurubripes. *Science*, *297*, 1301–1310.

Arendt, D., & Nubler-Jung, K. (1997). Dorsal or ventral: similarities in fate maps and gastrulation patterns in annelids, arthropods and chordates. *Mechanical Development*, *61*, 7–21.

Arendt, D., & Nübler-Jung, K. (1999). Rearranging gastrulation in the name of yolk: Evolution of gastrulation in yolk-rich amniote eggs. *Mechanical Development*, *81*, 3–22.

Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*, *408*, 796–815.

Baars, B. (1988). *A Cognitive Theory of Consciousness*. New York: Cambridge University Press.

Barde, Y. A. (1990). The nerve growth factor family. *Progress in Growth Factor Research*, *2*, 237–248.

Beddington, R. S., & Robertson, E. J. (1999). Axis development and early asymmetry in mammals. *Cell*, *96*, 195–209.

Benoit, P., & Changeux, J. P. (1975). Consequences of tenotomy on the evolution of multiinnervation in developing rat soleus muscle. *Brain Research*, *99*, 354–358.

Benoit, P., & Changeux, J. P. (1978). Consequences of blocking the nerve with a local anaesthetic on the evolution of multiinnervation at the regenerating neuromuscular junction of the rat. *Brain Research*, *149*, 89–96.

Bourgeois, J. P. (2001). Synaptogenesis in the neocortex of the newborn: the ultimate frontier for individuation? In H. Lagercrantz, M. Hanson, P. Evrard, & C. Rodeck (Eds.), *The Newborn Brain: Neuroscience and Clinical Applications* (pp. 91–113). Cambridge: Cambridge University Press.

Bourgeois, J. P. (2003). Le développement de la connectivité cérébrale: étape ultime de l'individuation? In J. P. Changeux (Ed.), *Gènes et culture* (93–116). Paris: Odile Jacob.

Brodmann, K. (1912). Neue Ergebnisse über die vergleichende histologische Lokalisation der grosshirnrinde Mitbesonderer Berücksichtigung des Stirnhirns. *Anatomische Anzeiger*, *41*, 157–216.

C. elegans Sequencing Consortium. (1998). Genome sequence of the nematode C. elegans: A platform for investigating biology. *Science*, *282*, 2012–2018. Review. Erratum in. *Science*, 283 (5398), 35; *Science 283* (5410), 2103; *Science 285* (5433), 1493.

Cabelli, R. J., Hohn, A., & Shatz, C. J. (1995). Inhibition of ocular dominance column formation by infusion of NT-4/5 or BDNF. *Science*, *267*, 1662–1666.

Cantalupo, C., & Hopkins, W. D. (2001). Asymmetric Broca's area in great apes. *Nature*, *414*, 505.

Carmena, A., Gisselbrecht, S., Harrison, J., Jimenez, F., & Michelson, A. M. (1998). Cell fates in the Drosophila embryonic mesoderm. *Gene Development*, *12*, 3910–3922.

Carmignoto, G., Canella, R., Candeo, P., Comelli, M. C., & Maffei, L. (1993). Effects of nerve growth factor on neuronal plasticity of the kitten visual cortex. *Journal of Physiology*, *464*, 343–360.

Carroll, S. B. (2003). Genetics and the making of Homo sapiens. *Nature*, *422*, 849–857.

Changeux, J. P. (1983). *L'Homme neuronal*. Paris: Fayard.

Changeux, J. P., Courrège, P., & Danchin, A. (1973). A theory of the epigenesis of neuronal networks by selective stabilization of synapses. *Proceedings of the National Academy of Sciences*, *70*, 2974–2978.

Changeux, J. P., & Mikoshiba, K. (1978). Genetic and "epigenetic" factors regulating synapse formation in vertebrate cerebellum and neuromuscular junction. *Progress in Brain Research*, *48*, 43–66.

Chervitz, S. A., Aravind, L., Sherlock, G., Ball, C. A., Koonin, E. V., Dwight, S. S., et al. (1998). Comparison of the complete protein sets of worm and yeast: orthology and divergence. *Science*, *282*, 2022–2028.

Crepel, F., Mariani, J., & Delhaye-Bouchaud, N. (1976). Evidence for a multiple innervation of Purkinje cells by climbing fibers in the immature rat cerebellum. *Journal of Neurobiology*, *7*, 567–578.

Davidson, E. H., Rast, J. P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C. H., et al. (2002). A genomic regulatory network for development. *Science*, *295*, 1669–1678.

de Kerchove d'Exaerde, A., Cartaud, J., Ravel-Chapuis, A., Seroz, T., Pasteau, F., Angus, L. M., et al. (2002). Expression of mutant Ets protein at the neuromuscular synapse causes alterations in morphology and gene expression. EMBO Rep., *3*, 1075–1081.

Deacon, T. (1997). *The Symbolic Species*. New York: Norton.

Dehaene, S., & Changeux, J. P. (2000). Reward-dependent learning in neuronal networks for planning and decision making. *Progress in Brain Research*, *126*, 217–229.

Dehaene, S., Kerszberg, M., & Changeux, J. P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences*, *95*, 14529–14534.

Dehaene, S., Sergent, C., & Changeux, J. P. (2003). A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proceedings of the National Academy of Sciences*, *100*(14), 8520–8525.

Driever, W., & Nüsslein-Volhard, C. (1988). The bicoid protein determines position in the Drosophila embryo in a concentration-dependent manner. *Cell*, *54*, 95–104.

Eckert, M. A., Leonard, C. M., Molloy, E. A., Blumenthal, J., Zijdenbos, A., & Giedd, J. N. (2002). The epigenesis of planum temporale asymmetry in twins. *Cerebral Cortex*, *12*, 749–755.

Edelman, G. M. (1988). Morphoregulatory molecules. *Biochemistry*, *27*(10), 3533–3543.

Elliott, T., & Shadbolt, N. R. (1998). Competition for neurotrophic factors: mathematical analysis. *Neural Computation*, *10*, 1939–1981.

Elston, G. N. (2000). Pyramidal cells of the frontal lobe: All the more spinous to think with. *Journal of Neuroscience*, *20*, RC95.

Elston, G. N. (2003). Cortex, cognition and the cell: New insights into the pyramidal neuron and prefrontal function. *Cerebral Cortex*, *13*, 1124–1138.

Enard, W., Przeworski, M., Fisher, S. E., Lai, C. S., Wiebe, V., Kitano, T., Monaco, A. P., & Paabo, S. (2002). Molecular evolution of FOXP2, a gene involved in speech and language. *Nature*, *418*, 869–872.

Gannon, P. J., Holloway, R. L., Broadfield, D. C., & Braun, A. R. (1998). Asymmetry of chimpanzee planum temporale: Humanlike pattern of Wernicke's brain language area homolog. *Science*, *279*, 220–222.

Gilad, Y., Bustamante, C. D., Lancet, D., & Paabo, S. (2003). Natural selection on the olfactory receptor gene family in humans and chimpanzees. *American Journal of Human Genetics*, *73*, 489–501.

Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., et al. (1996). Life with 6,000 genes. *Science*, *274*(5287), 546, 563–567.

Gouze, J. L., Lasry, J. M., & Changeux, J. P. (1983). Selective stabilization of muscle innervation during development: A mathematical model. *Biological Cybernetics*, *46*, 207–215.

Halfon, M. S., Carmena, A., Gisselbrecht, S., Sackerson, C. M., Jimenez, F., Baylies, M. K., & Michelson, A. M. (2000). Ras pathway specificity is determined by the integration of multiple signal-activated and tissue-restricted transcription factors. *Cell*, *103*, 63–74.

Hoch, M., Seifert, E., & Jackle, H. (1991). Gene expression mediated by cis-acting sequences of the Kruppel gene in response to the Drosophila morphogens bicoid and hunchback. *EMBO Journal*, *10*, 2267–2278.

Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Computational Neurology*, *387*, 167–178.

International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*, 860–921. Erratum in Nature *412* (6846), 565; *Nature 411* (6838), 720.

Kano, M., Hashimoto, K., Kurihara, H., Watanabe, M., Inoue, Y., Aiba, A., & Tonegawa, S. (1997). Persistent multiple climbing fiber innervation of cerebellar Purkinje cells in mice lacking mGluR1. *Neuron*, *18*, 71–79.

Katz, L. C., & Shatz, C. J. (1996). Synaptic activity and the construction of cortical circuits. *Science*, *274*, 1133–1138.

Kerszberg, M. (1996). Accurate reading of morphogen concentrations by nuclear receptors: A formal model of complex transduction pathways. *Journal of Theoretical Biolology*, *183*, 95–104.

Kerszberg, M., & Changeux, J. P. (1994a). A model for reading morphogenetic gradients: autocatalysis and competition at the gene level. *Proceedings of the National Academy of Sciences*, *91*, 5823–5827.

Kerszberg, M., & Changeux, J. P. (1994b). Partners make patterns in morphogenesis. *Current Biology*, *4*, 1046–1047.

Kerszberg, M., & Changeux, J. P. (1998). A simple molecular model of neurulation. *Bioessays*, *20*, 758–770.

Kerszberg, M., Dehaene S., & Changeux J. P. (1992). Stabilization of complex input-outpout functions in neural clusters formed by synapse selection. *Neural Networks*, *5*, 403–413.

Lai, C., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F., & Monaco, A. P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*, *413*, 519–523.

Lai, C. S., Gerrelli, D., Monaco, A. P., Fisher, S. E., & Copp, A. J. (2003). FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain*, *126*, 2455–2462.

Levi-Montalcini, R. (1987). The nerve growth factor 35 years later. *Science*, *237*, 1154–1162.

Levinthal, F., Macagno, E., & Levinthal, C. (1976). Anatomy and development of identified cells in isogenic organisms. *Cold Spring Harbor Symposin in Quantitative Biology*, *40*, 321–331.

Lohof, A. M., Ip, N. Y., & Poo, M. M. (1993). Potentiation of developing neuromuscular synapses by the neurotrophins NT-3 and BDNF. *Nature*, *363*, 350–353.

Maffei, L., Berardi, N., Domenici, L., Parisi, V., & Pizzorusso, T. (1992). Nerve growth factor (NGF) prevents the shift in ocular dominance distribution of visual cortical neurons in monocularly deprived rats. *Journal of Neuroscience*, *12*, 4651–4662.

Mannervik, M., & Levine, M. (1999). The Rpd3 histone deacetylase is required for segmentation of the Drosophila embryo. *Proceedings of the National Academy of Sciences*, *96*, 6797–6801.

Mannervik, M., Nibu, Y., Zhang, H., & Levine, M. (1999). Transcriptional coregulators in development. *Science*, *284*, 606–609.

Meinhardt, H., & Gierer, A. (1974). Applications of a theory of biological pattern formation based on lateral inhibition. *Journal of Cell Science*, *15*, 321–346.

Mountcastle, V. (1998). *Perceptual neuroscience: The cerebral cortex*. Cambridge, MA: Harvard University Press.

Mouse Genomic Sequencing Consortium. (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature*, *420*, 520–562.

Nguyen, Q. T., Parsadanian, A. S., Snider, W. D., & Lichtman, J. W. (1998). Hyperinnervation of neuromuscular junctions caused by GDNF overexpression in muscle. *Science*, *279*, 1725–1729.

O'Donovan, C., Apweiler, R., & Bairoch, A. (2001). The human proteomics initiative (HPI). *Trends in Biotechnology*, *19*, 178–181.

Olesen, P. J., Nagy, Z., Westerberg, H., & Klingberg, T. (2003). Combined analysis of DTI and fMRI data reveals a joint maturation of white and grey matter in a fronto-parietal network. *Brain Research/Cognitive Brain Research*, *18*, 48–57.

Purves, D., & Lichtman, J. W. (1980). Elimination of synapses in the developing nervous system. *Science*, *210*, 153–157.

Rakic, P. (1988). Specification of cerebral cortical areas. *Science*, *241*, 170–176.

Rauschecker, J. P., Tian, B., & Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science*, *268*, 111–114.

Rockel, A. J., Hiorns, R. W., & Powell, T. P. (1980). The basic uniformity in structure of the neocortex. *Brain*, *103*, 221–244.

Rubin, G. M., Yandell, M. D., Wortman, J. R., Gabor Miklos, G. L., Nelson, C. R., Hariharan, I. K., et al. (2000). Comparative genomics of the eukaryotes. *Science*, *287*, 2204–2215.

Sanes, J. R., & Lichtman, J. W. (1999). Development of the vertebrate neuromuscular junction. *Annual Reviews of Neuroscience*, *22*, 389–442.

Schaeffer, L., de Kerchove d'Exaerde, A., & Changeux, J. P. (2001). Targeting transcription to the neuromuscular synapse. *Neuron*, *31*, 15–22.

Schaeffer, L., Duclert, N., Huchet-Dymanus, M., & Changeux, J. P. (1998). Implication of a multisubunit Ets-related transcription factor in synaptic expression of the nicotinic acetylcholine receptor. *EMBO J.*, *17*, 3078–3090.

Semendeferi, K., Armstrong, E., Schleicher, A., Zilles, K., & Van Hoesen, G. W. (2001). Prefrontal cortex in humans and apes: A comparative study of area 10. *American Journal of Physical Anthropology*, *14*, 224–241.

Semendeferi, K., Damasio, H., Frank, R., & Van Hoesen, G. W. (1997). The evolution of the frontal lobes: A volumetric analysis based on three-dimensional reconstructions of magnetic resonance scans of human and ape brains. *Journal of Human Evolution*, *32*, 375–388.

Semendeferi, K., Lu, A., Schenker, N., & Damasio, H. (2002). Humans and great apes share a large frontal cortex. *Nature Neuroscience*, *5*, 272–276.

Smolen, P., Baxter, D. A., & Byrne, J. H. (2000). Mathematical modeling of gene networks. *Neuron*, *26*, 567–580.

Snyder, M., & Gerstein, M. (2003). Genomics: Defining genes in the genomics era. *Science*, *300*, 258–260.

Sretavan, D., & Shatz, C. J. (1986). Prenatal development of cat retinogeniculate axon arbors in the absence of binocular interactions. *Journal of Neuroscience*, *6*, 990–1003.

Steinmetz, H., Herzog, A., Schlaug, G., Huang, Y., & Jancke, L. (1995). Brain (A) symmetry in monozygotic twins. *Cerebral Cortex*, *5*, 296–300.

Thoenen, H. (1995). Neurotrophins and neuronal plasticity. *Science*, *270*, 593–598.

Venter, J. C., et al. (2001). The sequence of the human genome. *Science*, *291*, 1304–1351. Erratum in *Science 292* (5523), 1838.

Von Economo, C. (1929). *The Cytoarchitectonics of the Human Cerebral Cortex*. Oxford: Oxford University Press.

Wiesel, T. N., & Hubel, D. H. (1963). Effects of visual deprivation on morphology and physiology of cells in the cats lateral geniculate body. *Journal of Neurophysiology*, *26*, 978–993.

II Putative Prerogatives of the Human Brain and Their Evolutionary Precursors

6 Quantitative Thinking: From Monkey to Human and Human Infant to Human Adult

Elizabeth M. Brannon

Is thought possible in the absence of language? How much of human thought is defined by our capacity for language? If thought is critically dependent on language, then are animals and preverbal human infants bereft of thought? These questions have a long history and have captivated philosophers, biologists, and psychologists alike. One view, championed early on by Descartes (1646/1970) and later Locke (1690/1975) was that animals are incapable of thought primarily because they lack language—the key ingredient for the formation of ideas. For Descartes, famous for his belief in the separation of mind and body, both animals and humans possessed body, open to scientific study, and only humans possessed the intangible mind—a soul that permits intelligent voluntary behavior and thought. Although Descartes noted some fundamental similarities between animals and humans in the area of emotion and memory (1649/1989, p. 48), it was not until two centuries later that Darwin revolutionized the study of behavior by theorizing that there was evolutionary continuity between the animal and human mind. Darwin (1871/1920, p. 128) wrote, "the difference in mind between man and the higher animals, great as it is, certainly is one of degree and not of kind." This perspective leads to the prediction that despite lacking language, precursors of many complex human cognitive capacities should be present in nonhuman animals. Support for intellectual continuity between animals and humans comes from the results of a century of controlled experiments on animal cognition that have revealed limited precursors to hallmarks of humanity such as language, concept formation, cooperation (chapter 9, this volume), analogical reasoning, theory of mind, music appreciation, and mathematics throughout the animal kingdom.

Similarly, since the adult mind somehow emerges from the immature mind, the seeds of adult cognition must be present at some point in childhood or infancy before they are instantiated in their adult form. Nevertheless, there is great controversy over how early in development we see precursors to adult human cognition and whether developmental change is continuous, reflecting only quantitative change, or sometimes discontinuous reflecting qualitative changes. Some perspectives view the acquisition of language as a force that radically changes thought, while others view language

as playing a more marginal role in development. Although the emergence of language may fundamentally change the nature of thought, a plethora of studies demonstrate that infants represent complex aspects of their environment well before productive language emerges. For example, contra to Piaget, infants as young as 3.5 months of age represent hidden objects (Baillargeon, 1987), infants as young as 6 months of age form categories (e.g., Mareschal & Quinn, 2001), and infants as young as 8 weeks of age remember a mobile that moves contingently with their own movements 2 weeks later (Rovee-Collier, 1999). Thus, while there is no doubt that the mind of the infant experiences vast changes from the helpless newborn with 20/600 vision, dominated by a suite of reflexes, endowed with a head 25 percent of its mature size, and completely dependent on its caregivers for warmth, sustenance, and mobility, there is also strong evidence that many of the building blocks of human cognition are present in early infancy.

Studies of animal and infant cognition show us that thought is possible without language. Our challenge now is to describe the nature of languageless thought. By investigating cognition in animals and infants we can begin to identify the primitives that form the foundations of adult human cognition.

The Case of Number

The study of how the mind represents number is a paradigmatic example of complex cognition that can and does take place in the absence of language. As such, number is a good case study for studying the nature of thought in the absence of language and has brought together disparate fields such as developmental psychology, animal cognition, and cognitive neuroscience.

The study of animal number concepts got off to a tumultuous beginning. In the early twentieth century it was claimed that a horse named Clever Hans could count and even do complex mathematics (Rosenthal, 1911/1965). When presented with a mathematical problem on a chalkboard, Hans would tap out the answer with his hoof. Unbeknownst to the horse trainer, Hans was actually determining when to cease tapping his hoof by attending to subtle and unconscious cues that the human questioners emitted. When the humans were blind to the mathematical question, the horse failed the test! This blight on the study of animal numerical ability resulted in a healthy skepticism toward animal mathematics and led to rigorous controls to eliminate all potential nonnumerical cues.

Over the last century of investigation evidence that animals can represent number and manipulate their representations of number has accumulated (for a review, see Brannon & Roitman, 2003). The main finding from these varied paradigms and approaches has been that animal number discrimination follows Weber's law which states that the change in intensity required to notice a difference between two stimuli

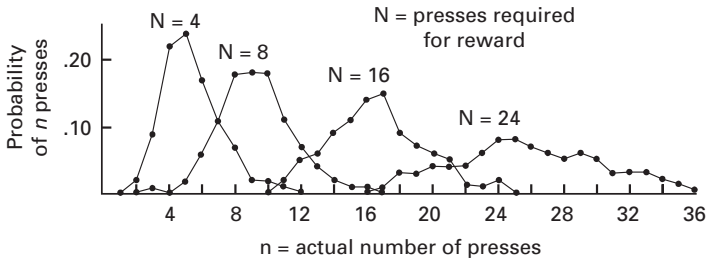


Figure 6.1
The probability of signaling response completion as a function of the number of responses the rat made and the number that was required to obtain reward obtained by Platt & Johnson (1971).

is equal to a constant proportion of the standard stimulus. In other words, the ratio of two stimuli rather than their absolute difference determines whether they can be discriminated from each other. This is illustrated nicely by a paradigm first designed by Mechner (1958) and later adapted by Platt and Johnson (1971). Rats were required to signal when they had completed *n* lever presses by poking their nose into a hole equipped with a photoelectric sensor. Figure 6.1 shows that the number of responses the rats made before head poking was roughly normally distributed around the required number. Thus when required to make *x* presses, the animals were much more likely to make $x - 1$ or $x + 1$ presses than $x - 4$ or $x + 4$ presses. Furthermore, the standard deviation of the distribution of the obtained number of responses increased linearly with the required number of responses. Such results suggest that rats represent number not as precise values, but instead as fuzzy magnitudes that overlap more and more with increasing size (see Dehaene, Dehaene-Lambertz, & Cohen, 1998; Gallistel & Gelman, 2000).

The animal research tells us that number representation is not dependent on language and leads to the question of whether preverbal human infants also have a nonverbal system for representing number. Is it possible that animals and infants might share a phylogenetically ancient and ontogenetically conservative mechanism for representing number? A handful of early studies using infants' looking behavior as a dependent measure, suggested that human infants could discriminate small numerical values such as 2 versus 3 but could not discriminate larger values with the same ratio such as 4 versus 6 (e.g., Starkey & Cooper, 1980). Such findings suggested that infants might be using a mechanism that was limited to representing small numerical values, and prompted many to believe that infants were "subitizing," a term used in the adult literature to describe a rapid perceptual appreciation of small numbers that was thought to occur by a parallel rather than iterative process (Kaufman, Lord, Reese, & Volkmann, 1949). In conflict with this conclusion, it has recently been found

that young infants can represent large values in an approximate manner. In the first study Xu and Spelke (2000) found that young infants habituated to 16 dots looked longer at test displays that contained 8 dots and vice versa, even when surface area and density were strictly controlled. Importantly, while 6-month old infants could discriminate 8 versus 16, they could not discriminate 8 versus 12. And more recent research using auditory stimuli suggests that the precision of these numerical representations increases with development so that by 9 months of age infants can discriminate large values with a 2:3 ratio (Lipton & Spelke, 2003). The finding that infants can discriminate large values if they are sufficiently disparate in numerical magnitude suggests that Weber's law controls numerical discrimination in infancy and that infants, like nonhuman animals, are representing number as continuous mental magnitudes.

However, other findings with human infants provide a more complicated story. First, a handful of studies have tested infants' ability to discriminate small values using stimulus controls for surface area and other continuous variables that were largely absent in earlier studies, and found that infants do not represent small numerical values independently of continuous variables (e.g., Clearfield & Mix, 1999; Feigenson, Carey, & Spelke, 2002; Xu, 2003). Some interpret these results to mean that infants are not doing anything numerical at all and that all positive results are a byproduct of a lack of control for continuous dimensions (e.g., Mix, Huttenlocher, & Levine, 2002). Others argue that infants have a repertoire of mechanisms and use them in different contexts. Although one mechanism represents number as analog magnitudes, a second mechanism may represent number almost by accident and be more sensitive to continuous variables. It is also likely that infants possess additional mechanisms that keep track of continuous variables with no regard for number (e.g., Clearfield & Mix, 1999).

This second system which functions to represent number, but does so almost by accident, has been termed the object-file system and is thought to be limited in its capacity and represent the numerosity of small sets via symbols that represent each individual object (see Kahneman, Treisman, & Gibbs, 1992). Several researchers have offered some version of this hypothesis (e.g., Feigenson, Carey, & Spelke, 2002; Koechlin, Dehaene, & Mehler, 1997; Scholl & Pylyshyn, 1999; Uller, Carey, Huntley-Fenner, & Klatt, 1999; Simon, 1997; Spelke, 2000; Xu, 2003). The idea is that a pointer is assigned to each item in a visual array and that there is a limited number of pointers available (Pylyshyn, 2001). In contrast to the magnitude system, which results in a symbol that is isomorophic with the numerosity of the set it serves to represent, there is no symbol in the object-file system that represents the set of objects. Instead, the object-file system results in symbols that represent each individual object. Supporting the claim that infants use object-files that function to represent number, a handful of studies have shown that infants succeed at discriminating small values such

as 1 versus 2, and 2 versus 3 but fail at discriminating sets with larger values even when the sets involve favorable ratios such as 2 versus 4, or 4 versus 8 (Feigenson, Carey, & Hauser, 2002; Feigenson & Carey, 2003). On the flip side, when continuous variables are strictly controlled infants succeed at representing large values such as 4 versus 8 and fail at representing small values such as 2 versus 4 (Xu, 2003).

One possible explanation for this pattern of results is that some contexts activate an object-file mechanism and do not elicit the formation of analog magnitudes, in these cases infants will fail to detect differences between large numerical sets (e.g., Feigenson, Carey, & Hauser, 2002). Conversely, other situations elicit the formation of mental magnitudes and do not activate the object-file mechanism, in these contexts infants can represent large numerical sets and in some may even fail to represent small sets (e.g., Xu, 2003). This analysis leaves open the puzzling question of why infants do not form analog magnitudes when presented with small sets. Another question is whether the object-file system, like the analog magnitude system, is phylogenetically ancient. In general, research with animals has not revealed the same dissociation between the object-file and analog magnitude systems. For example, many tasks require that animals discriminate small and large values and find graded levels of responding for intermediate values with small and large values being represented as analog magnitudes (e.g., Meck & Church, 1983). In a recent study in my laboratory, three monkeys were trained in a conceptual numerical delayed match-to-sample task where they were presented with samples that contained 2 or 8 elements and then given a choice between two stimuli that contained 2 or 8 elements but differed from the sample in size, shape, and color of the elements (Jordan & Brannon, 2003). Example stimuli are shown in the inset of figure 6.2. After the monkeys mastered the 2 versus 8 task, they were then tested with samples of all intermediate values and the values 1 and 9 and again required to choose between stimuli with 2 or 8 elements. Only trials with 2 or 8 as the sample were reinforced and these occurred at a high frequency to prevent extinction. Figure 6.2 shows that the probability of choosing the stimulus with eight as opposed to two elements increased systematically with the number of elements in the display. In addition the data suggest that the monkeys viewed stimuli with four elements as equally similar to two and eight (i.e., the point of subjective equality was at the geometric mean). These data illustrate that rhesus monkeys' numerical similarity judgments are graded throughout the range 1–9 and provide no evidence of discontinuities or an inability to represent small values as analog magnitudes. However, in some studies Hauser and colleagues have found that monkeys fail to differentiate large sets even when they can discriminate small sets that differ by the same ratio (Hauser, Carey, & Hauser, 2000). Future work will need to determine whether monkeys, like human infants, use two distinct systems for representing number and whether the conditions that elicit these systems differ for infants and monkeys.

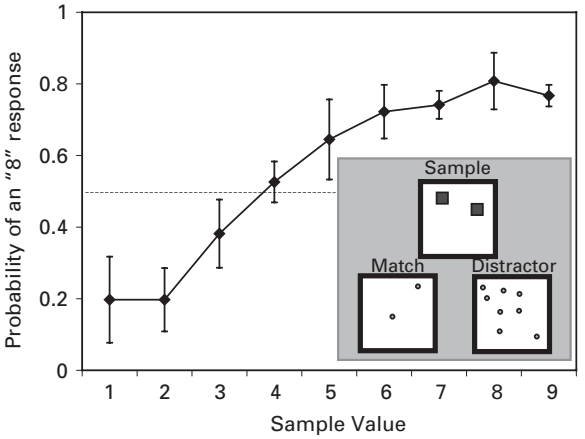


Figure 6.2
The probability with which an exemplar of the numerosity eight is chosen rather than an exemplar of the numerosity two, plotted as a function of the numerosity of the sample. Data reflect an average of 3 monkeys over approximately 100 unreinforced trials for each sample numerosity. (*Inset*) Example trial for a delayed conceptual numerical match-to-sample task. In this example, the correct choice and distractor have elements that are identical in color size and shape and that differ from the elements in the sample on all three dimensions.

Arithmetic in Animals and Infants

Beyond representing number, an important question is whether animals and infants actually manipulate their numerical representations in arithmetic operations. One of the main advantages to representing number as continuous mental magnitudes is that information about ordinal relations is implicit in the representations and arithmetic such as addition and subtraction is straightforward (Gallistel & Gelman, 2000). Do nonhuman animals appreciate the ordinal relations between numerosities? Brannon and Terrace (1998, 2000) have addressed this question in a series of experiments. Rhesus monkeys were presented with stimuli on a touch-sensitive screen and required to respond in ascending numerical order. On each trial, four stimuli were presented in a random spatial configuration with each stimulus containing 1, 2, 3, or 4 elements. The elements were either simple geometric shapes or more complex clipart shapes, and the stimuli contained a homogeneous or heterogeneous collection of elements. Across 35 training and 150 test sets, nonnumerical cues, such as surface area, were randomly varied so that number was the only valid cue as to the ordinal position of each stimulus in the four-item sequence (figure 6.3a).

After the monkeys learned each of 35 training sets to a performance criterion, they were tested with 150 novel stimulus sets, each presented only for a single trial. These

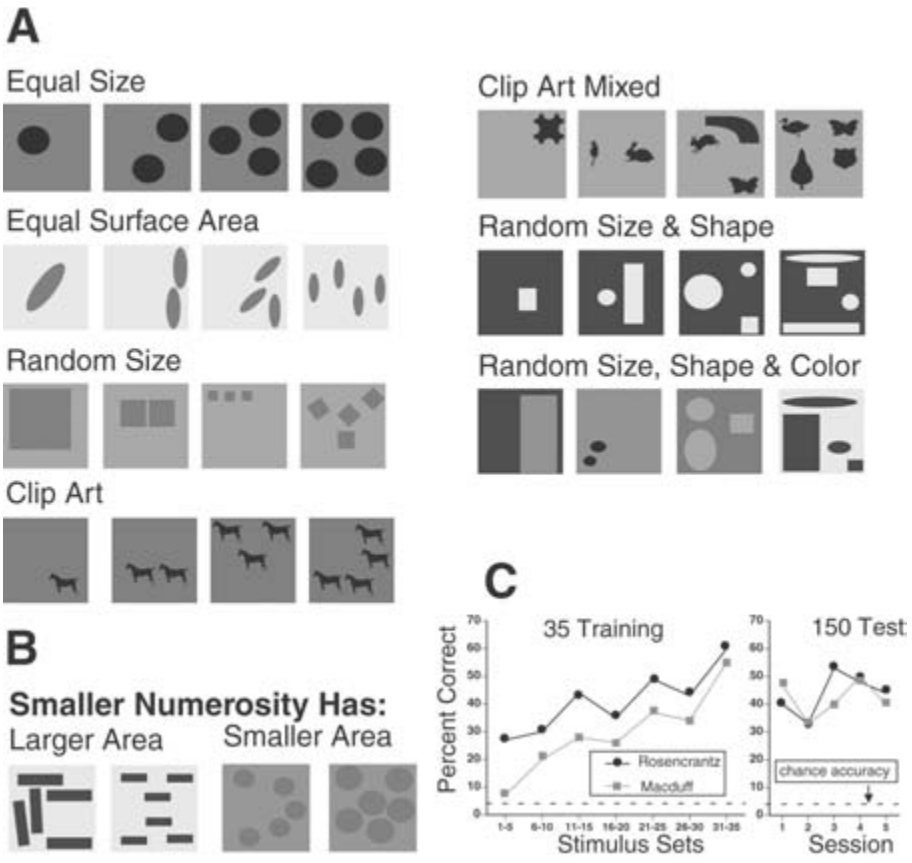


Figure 6.3

(*A*) Exemplars of the seven different types of stimulus sets used by Brannon & Terrace (1998). *Equal size*: elements were of same size and shape. *Equal area*: cumulative area of elements was equal. *Random size*: element size varied randomly across stimuli. *Clip art*: identical nongeometric elements selected from clip art software. *Clip art mixed*: clip art elements of variable shape. *Random size and shape*: elements within a stimulus were varied randomly in size and shape. *Random size, shape and color*: same as previous with background and foreground colors varied between stimuli. (*B*) Examples of stimulus sets used in the pair-wise numerosity test. (*C*, left): Percent correct for 35 training sets. Each was presented for 60 trials, and each data point reflects the average of 5 sessions (300 trials). (*Right*): Percent correct for 150 trial unique test sets tested in 5 test sessions with 30 trials each. Chance accuracy is less than 4 percent in this task (.25 * .33 * .5). (Reprinted from Brannon & Terrace, 1998.)

test sessions provided no opportunity to memorize specific stimulus features; thus, above chance performance would be evidence that the monkeys used a numerical rule. Figure 6.3c shows that the monkeys' performance improved rapidly over the 35 training sets, and their performance was not impaired in the five test sessions that were composed of trial-unique stimulus sets. In addition, monkeys performed above chance on all seven different stimulus classes. These data demonstrate that rhesus monkeys can discriminate the numerosities 1–4 without using nonnumerical cues such as shape, color, or element size or cumulative element surface area.

In a second experiment, Brannon and Terrace (1998, 2000) tested whether the monkeys appreciated the ordinal relations between the numerosities or instead represented the numerosities categorically. The same two monkeys were tested on their ability to order pairs of the numerosities 1–9 after the 1–4 training. The critical question was whether the monkeys would be able to extrapolate what to do with the novel numerosities 5–9 from their training with the numerosities 1–4. The monkeys were presented with all the possible pairs of the numerosities 1–9 and were expected to respond to the smaller value first. The smaller number had a larger cumulative surface area than the larger number on half of the trials (see figure 6.3b). To provide a pure test of ordinal numerical knowledge, the monkeys were not reinforced on any trial that contained a novel numerical value. Thus, only trials that contained two exemplars of the numerosities 1–4 were reinforced. The other 30 pairs were tested in the absence of positive or negative reinforcement. This was a powerful test of ordinal numerical knowledge because there was no laboratory-learned basis by which the monkeys could judge the ordinal relations between numerical values that were outside the training range. For example, if one learned only the beginning of a new alphabet there would be no basis for ordering the latter part.

The monkeys' performance was extremely good for pairs composed of two familiar numerosities (e.g., 1 vs. 3, or 2 vs. 4) and pairs composed of 1 familiar and 1 novel value (e.g., 2 vs. 8, or 3 vs. 6). Most important, however, the monkeys performed above chance expectations on pairs composed of two novel values (e.g., 6 vs. 8). These results indicate that monkeys represent the ordinal relations between numerosities and do so spontaneously even when they could have instead formed arbitrary numerical categories and learned an arbitrary ordering of these nominal categories. The same pattern of results has since been obtained with a squirrel monkey and a baboon (Smith, Piel, & Candland, 2003).

A second piece of data from the same series of studies by Brannon and Terrace provides further evidence that monkeys represent the ordinal relations between numerosities. Brannon and Terrace originally attempted to train one of the monkeys to respond to the numerosities 1–4 in an arbitrary nonmonotonic order (Brannon & Terrace, 2000). Figure 6.4 shows that despite extended training on 13 different sets of stimuli the monkey never learned to respond in the order 3-1-4-2. Subsequently when given

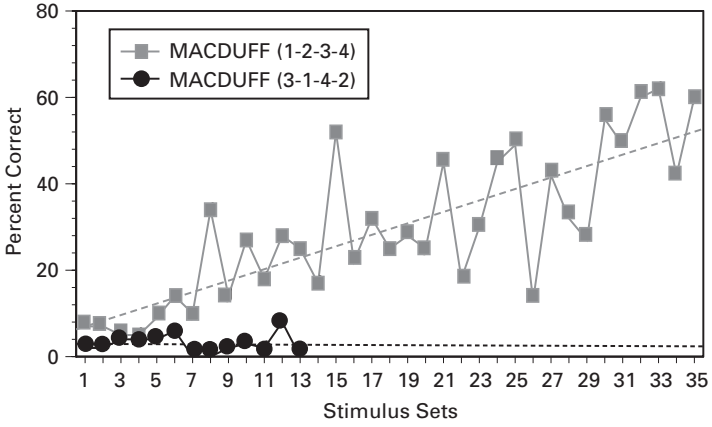


Figure 6.4
Performance for a monkey required to respond in the order 3-1-4-2 to visual arrays on a touch sensitive screen (black circles) and subsequently required to respond to the same stimuli in ascending numerical order (gray squares). (Reprinted from Brannon & Terrace, 2000.)

new stimulus sets and required to respond in ascending order (1-2-3-4) the monkey's performance quickly accelerated. These data suggest that the monkey's inherent ordinal representation of the numerosities 1–4 prevented it from responding in an arbitrary nonmonotonic order. Together these two findings suggest that number is a meaningful stimulus dimension for rhesus monkeys.

The experiments reviewed above suggest that rhesus monkeys represent ordinal relations between numerosities, what about human infants? Early research suggested that infants do not represent ordinal relationships between numerosities until they are in the second year of life (Cooper, 1984; Strauss & Curtis, 1984). However, Brannon (2002) recently found evidence that, infants as young as 11 months of age, appreciate ordinal numerical relations. In that study, infants were habituated to ascending or descending numerical sequences. The sequences consisted of three visual exemplars of numerosities. Between trials the absolute values changed (1-2-4, 2-4-8, or 4-8-16) but the ordinal direction was constant. Infants were then tested with the new numerical values, 3-6-12 where the ordinal direction was maintained or reversed. As shown in figure 6.5, the size of the elements varied such that cumulative surface area did not consistently increase or decrease with number and density was constant across the three test values. Eleven-month-old infants looked for significantly longer when the ordinal direction was reversed (figure 6.6).

These results suggest that by 11 months of age preverbal infants appreciate the ordinal relations between numerosities. However, an alternative explanation of the Brannon (2002) finding is that infants may have simply noticed the larger absolute

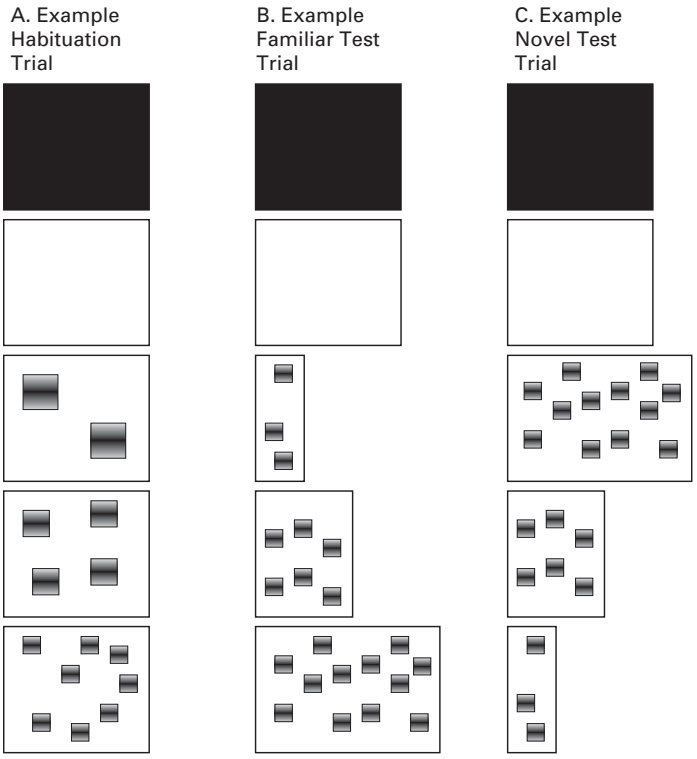


Figure 6.5
The five frames of a sample (*A*) habituation, (*B*) familiar test, and (*C*) novel test trial. The dotted lines surrounding each stimulus were not visible in the experiment. The sequence repeated, beginning with the black screen, until specific criteria were met.

difference between the average value of the first numerosity in the habituation sequences and the first value in the novel test sequences. For example infants habituated to the ascending order viewed a first stimulus with an average value of 2.33 and were then tested with 3-6-12 or 12-6-3. Infants habituated to the descending order viewed a first stimulus with an average value of 9.33 and were also tested with 3-6-12 or 12-6-3. In both cases, the sequence of the novel ordinal direction had a much larger difference in absolute value of the first stimulus (i.e., 2.33 differs more from 12 than from 3 and 9.33 differs more from 3 than from 12). However against this interpretation, Brannon (2003) found that 11-month-old infants tested in the same experimental design but with 2-item sequences that were missing the third stimulus of each sequence, did not look longer at the novel ordinal direction. This suggests that infants could not have been merely attending to the change in absolute value of the first

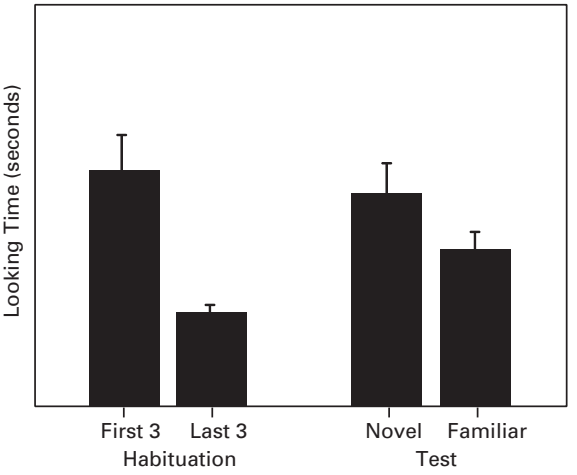


Figure *6.6*
Mean looking time (± SE) into the first three and last three habituation trials and to novel and familiar test trials for 11-month-old infants.

numerical stimulus in habituation and test sequences. Instead by 11 months of age infants represent the ordinal relations between numerosities and may require at least three numerosities in a sequence for ordinal numerical relations to be salient.

A handful of studies have also addressed whether animals and human infants are capable of some form of addition and subtraction. In a now classic study, Wynn (1992) tested infants in a violation of expectancy paradigm where five-month-old infants viewed an addition or subtraction event followed by a possible or impossible outcome. The idea is that if infants keep track of the number of toys they see being placed behind a screen they should look longer when the screen is lowered to reveal an outcome that violates their expectations. For example, in an addition event, infants viewed a stage with a Mickey Mouse doll, which was then obscured by a raised screen. A hand subsequently entered the display and placed a second Mickey Mouse doll behind the screen. The screen was then lowered to reveal the expected outcome of two dolls or the unexpected impossible outcome of one doll. Infants looked longer at the impossible outcome of 1 compared to the possible outcome of 2. Furthermore subsequent conditions revealed that infants were representing the precise number because infants who viewed a 1 + 1 event also looked longer at the outcome of 3 compared to 2. In the original studies continuous variables such as volume or surface area were not controlled so that infants might have attended to the number of dolls behind a screen or the amount of doll stuff. In fact when Feigenson, Carey, & Spelke (2002) presented infants with a 1 small + 1 small = 1 large or 2 large event, infants looked

longer at the 2 large event compared to the 1 large event suggesting that they had encoded total surface area rather than number prompting the authors to argue that infants represent each object as a stripped down object-file that preserves some information about object size.

Hauser et al. (1996) adapted Wynn's paradigm and found that like human infants, rhesus monkeys looked longer when an impossible number of eggplants were revealed after an addition or subtraction event. Furthermore, Flombaum, Junge, and Hauser (submitted) found that rhesus monkeys again looked longer at impossible outcomes when large numbers that differed by a 1:2 ratio were tested and when continuous variables were controlled. In one experiment, monkeys watched as four eggplants were added to a stage that already contained four hidden eggplants and then looked longer at the outcome of four eggplants compared to eight eggplants. In a second experiment, monkeys saw one medium eggplant added to three medium eggplants and subsequently looked longer at eight small eggplants compared to four large eggplants that were equated in total surface area. The large values used in the Flombaum et al. study preclude the possibility that monkeys used object-files. In addition, Flombaum et al. found that when monkeys viewed a 2 + 2 event they did not look longer at six compared to four eggplants, suggesting that they required a 1:2 ratio and could not discriminate a 2:3 ratio and providing support for the idea that they were using a system that represents number as analog magnitudes and is sensitive to Weber's law.

Many questions remain. Are infants in the Wynn experimental design tracking objects using object-files that preserve surface area of the objects as suggested by Feigenson, Carey, and Spelke (2002)? Perhaps certain conditions would allow infants in the Wynn task to use analog representations of number, ignore surface area and represent larger values as rhesus monkeys seem to do? Do monkeys like human infants, use an object-file system to represent small values? What are the conditions that elicit these two very different representational systems and do the conditions differ for infants and nonhuman primates?

Adult Humans: Further Evidence for a Shared Nonverbal Number System

Like the languageless beings described above, adult humans also represent number as nonverbal mental magnitudes. In a classic study, Moyer and Landauer (1971) showed that when adults were required to choose the larger of two Arabic numerals, accuracy increased and latency to respond decreased with increasing numerical disparity. Furthermore, when distance was held constant performance decreased with increasing numerical magnitude; this is referred to as the magnitude or size effect. In other words, both accuracy and latency were modulated by the ratio of the quantities that the numerals represented. This robust and highly replicable finding has been

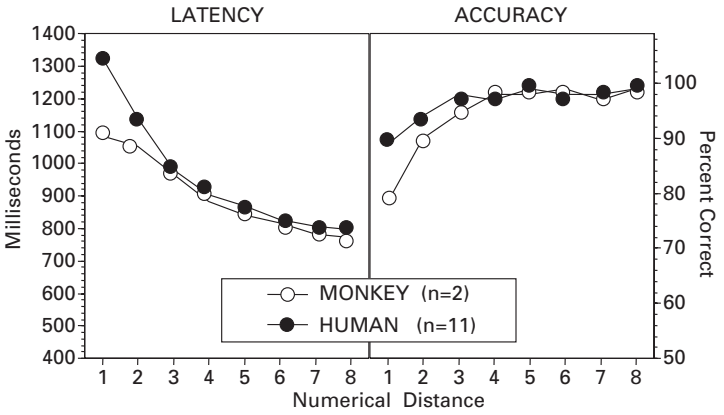


Figure 6.7
Accuracy and latency to the first response in a pair-wise numerical comparison task as a function of numerical disparity. Monkeys (light gray diamonds) and humans (dark gray squares) were required to respond first to the stimulus with the fewer number of elements.

interpreted to mean that Arabic numerals are represented as analog magnitudes, much like line length, brightness, or weight (see chapter 8, this volume).

The numerical distance effect can also be found in nonhuman animals. For example, Brannon and Terrace (2002) tested rhesus monkeys and college students in the same experiment, where both species were required to touch the smaller of two numerosities presented on a touch-screen. The stimuli were constructed such that the smaller numerosity had a larger cumulative surface area on half of the trials and all 36 possible pairings of the numerosities 1–9 were presented. Although monkeys worked for banana pellets whereas humans worked for course credit the tasks were otherwise identical. Figure 6.7 displays accuracy and latency to respond as a function of numerical disparity for each species and shows strikingly similar distance effects for the two species: both species were faster and more accurate as numerical disparity increased. Not shown here is the finding that when distance was held constant and size was increased, both species showed a tendency to decrease accuracy and increase reaction time. The similarity in the distance and size effects observed in monkeys and human adults provides strong support for the idea that animals and humans share a nonverbal system for representing number as mental magnitudes.

In another demonstration of the similarity between animal and human nonverbal number representations, Whalen, Gallistel, and Gelman (1999) tested human subjects in a modified version of the Platt and Johnson (1971) design described earlier. Adults were asked to make between 7 and 25 key presses as fast as they could without verbally counting. The results closely resembled the rat data obtained by Platt and

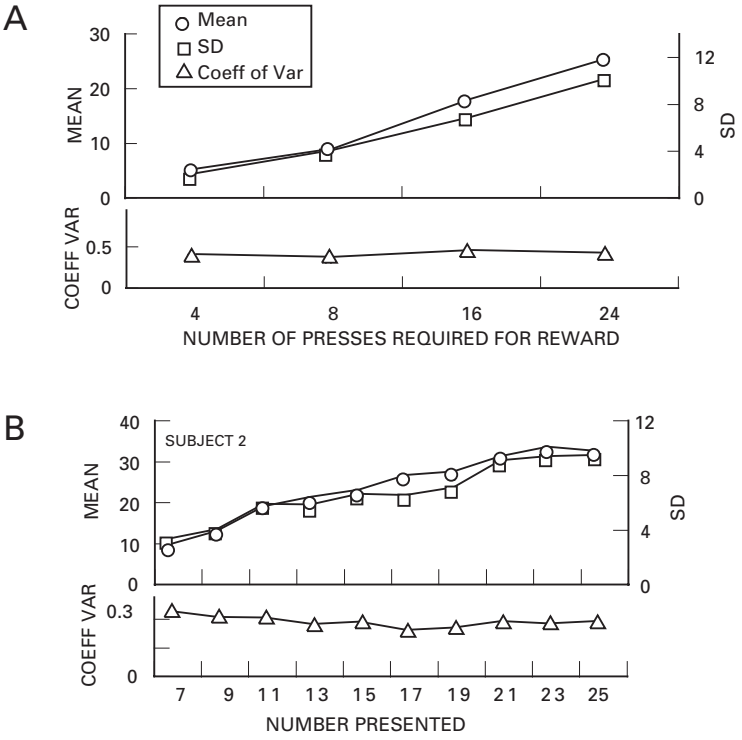


Figure 6.8

(*A*) The mean number of responses made (left axis, circles) and the standard deviation (right axis, squares) of the response distributions shown in (*A*), and the coefficient of variation (CV) which is the ratio of the standard deviation to the mean, as a function of the number required. (*B*) The mean (left axis, circles), standard deviation (right axis, squares), and CV (lower panel) as a function of the number of button responses required obtained by Whalen et al., (1999). The constant CV shown for rats in (*B*) and humans in (*C*) demonstrates that both species represent number with scalar variability. (Reprinted from Whalen et al., 1999.)

Johnson three decades earlier. Specifically, as shown in figure 6.8a and b, scalar variance was found for both species as indicated by the linear increase in the standard deviation in the response distributions as a function of the mean number of responses required and by the constant coefficient of variation obtained for both species (see also Cordes, Gallistel, & Gelman, 2001). These tasks appear to have tapped a nonverbal system for representing number in adult humans that is quite similar to that of rats! It was unlikely that humans were verbally counting since the standard deviation in the response distributions was proportional to the mean number of key presses subjects made and not proportional to the square root of the target count.

Further evidence that the subjects were not verbally counting is that in another experiment when subjects were instructed to verbally count, the variance was binomial and not scalar.

How Is Number Represented Nonverbally?

The data reviewed above suggests that animals, human infants, and human adults represent number as continuous mental magnitudes (See Gallistel & Gelman, 2000; Dehaene, Dehaene-Lambertz, & Cohen, 1998; Walsh, 2003). Strong evidence for this conclusion comes from studies that show that adults and monkeys show very similar distance and magnitude effects when comparing numerical dot displays (e.g., Brannon & Terrace, 2002) and studies that show that infants' numerical discrimination is well modulated by the ratio of the numerosities compared (Xu & Spelke, 2000; Lipton & Spelke, 2003). A separate question is how magnitude representations of number are constructed from sets of discrete entities. A few models have been proposed that result in analog magnitude representations of number. One model, termed the mode-control model (Meck & Church, 1983) or accumulator model (Gallistel & Gelman, 1992) holds that each discrete object or event results in the closing of a switch for a constant duration allowing pulses from a pacemaker to enter an accumulator. In this way a continuous magnitude is accumulated for any set of discrete elements that is isomorphic to the number it represents.

In contrast a model proposed by Dehaene and Changeux (1993) suggests that number is represented by a neural network that consists of three layers: an input "retina," a map of object locations, and an array of numerosity detectors. The map of object locations converts stimuli from the "retina" to a representation of each stimulus irrespective of object size. The location map sends its output to numerosity detectors, which consist of summation units and numerosity units. When the total activity from the output of the location map (which is proportional to numerosity) exceeds the summation unit's threshold, it will be activated. Finally, the summation clusters project to numerosity clusters, which represent the numerosities 1 through 5. A given numerosity cluster will be activated if the corresponding summation cluster is active, but those representing higher values are not. Therefore, presentation of stimuli with the same numerosity, despite differences in size, location and modality, results in the activation of the same numerosity detectors.

Two key differences between these models are that in contrast to Dehaene and Changeux's neural network model, the mode-control model predicts that time and number are represented with a single currency (see also Walsh, 2003) and that the process of forming a numerical representation is iterative and not parallel. Evidence that time and number are represented with a single representational currency comes from studies with rats and pigeons and is reviewed elsewhere (e.g., Brannon &

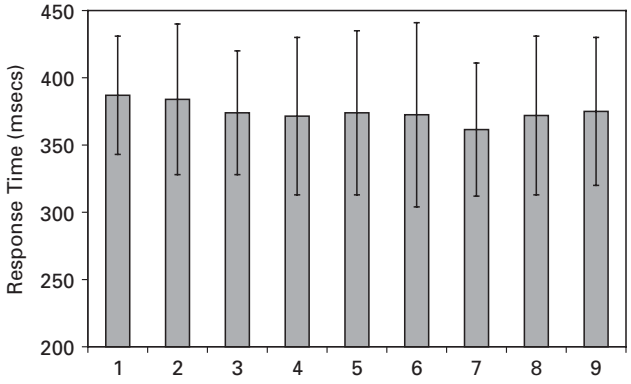


Figure 6.9
Reaction time to touch the sample as a function of sample numerosity in a delayed match-to-sample task. Data reflect an average of three monkeys.

Roitman, 2003). Very little evidence addresses the second question of whether animals or infants use an iterative or parallel mechanism when forming magnitude representations of number. Using the numerical conceptual MTS task described previously we have found suggestive evidence that the mechanism monkeys use in this task is not iterative. In this task the sample is presented until the monkey touches it. A touch eliminates the sample and produces the choice stimuli, thus it might be expected that monkeys would require longer to encode large values compared to small values. However, as shown in figure 6.9, we find that the RT to touch the sample does not vary with sample numerosity and is so fast it is almost at motor threshold. Providing additional support for a noniterative mechanism, Nieder and Miller (this volume) report that neurons in prefrontal and parietal cortex are selective for specific numerosities in the first 1–200 milliseconds of stimulus presentation.

Conclusions

It was once held that no thought could occur in the absence of language. This chapter, however, documents that both nonverbal animals and preverbal animals represent abstract number in the absence of language. Not addressed here is the entirely separate question of how language transforms the infant's representation of number (see Gelman & Cordes, 2001, and Carey, 2001, for opposing views). Instead, this chapter describes an emerging story whereby animals, human infants, and adults all share a nonverbal cognitive system for representing number as mental magnitudes that are an analog of number. This evolutionarily and developmentally primitive system results in fuzzy representations of number—where large values are confused more easily than

small values. Less clear is how these magnitude representations are formed. Does the nonverbal system follow the counting principles as suggested by Gallistel and Gelman (1992, 2000), or, alternatively, is the mechanism by which mental magnitudes are formed a parallel process? Future work should also examine the relationship between number representation and continuous stimulus variables such as time, space, and area (Walsh, 2003; see chapter 8, this volume) and elucidate whether these dimensions share a single representational currency.

Despite possession of an analog magnitude system, as evidenced by the ability to discriminate 8 versus 16 elements, the human infant sometimes fails to activate this system and instead represents small sets of objects via individual object-files. Is this object-file system the ontogenetic foundation for the adult preconceptual visual indexing system (e.g., Pylyshyn, 2001)? Is the object-file system like the analog magnitude system found throughout the animal kingdom? What are the contexts that activate one of these two systems and not the other? And do these contexts differ for nonhuman animals and human infants?

Thus a myriad of questions remain for future investigations of numerical cognition in human infants and animals. However, we can already say with some certainty that without language the minds of animals and human infants possess the seeds of quantitative thinking and a shared system for representing number as mental magnitudes.

Acknowledgments

The author thanks Stanislas Dehaene, Andreas Nieder, and Lisa Feigenson for comments on the chapter. Some of the research described in this paper was supported by NSF (ROLE and DLS) to E. M. Brannon, RO3 MH64955-01 to E. M. Brannon and G. R. Mangun, and a John Merck fellowship to E. M. Brannon.

References

Baillargeon, R. (1987). Object permanence in $3^1/_2$- and $4^1/_2$-month-old infants. *Developmental Psychology*, *23*(5), 655–664.

Brannon, E. M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition*, *83*, 223–240.

Brannon, E. M. (2003). "Numerical thinking: A comparative study." Paper presented at the Eastern Psychological Association, 2003.

Brannon, E. M., & Roitman, J. (2003). Nonverbal representations of time and number in nonhuman animals and human infants. In Warren Meck (Ed.), *Functional and Neural Mechanisms of Interval Timing* (pp. 143–182). New York, NY: CRC Press.

Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1–9 by monkeys. *Science*, *282*, 746–749.

Brannon, E. M., & Terrace, H. S. (2000). Representation of the numerosities 1–9 by rhesus monkeys (Macaca mulatta). *Journal of Experimental Psychology*: *Animal Behavior Processes*, *26*, 31–49.

Brannon, E. M., & Terrace, H. S. (2002). The evolution and ontogeny of ordinal numerical ability. In M. Bekoff, C. Allen, & G. M. Burghardt (Eds.), *The Cognitive Animal* (pp. 197–204). Cambridge, MA: MIT Press.

Carey, S. (2001). On the very possibility of discontinuities in conceptual development. In E. Dupoux (Ed.), *Language, Brain, and Cognitive Development: Essays in Honor of Jacques Mehler* (pp. 303–324). Cambridge, MA: MIT Press.

Clearfield, M. W., & Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science*, *10*, 408–411.

Cooper, R. G. J. (1984). Early number development: Discovering number space and addition and subtraction. In C. Sophian (Ed.), *Origins of Cognitive Skills* (pp. 157–192). Hillsdale, NJ: Erlbaum.

Cordes, S., Gelman, R., Gallistel, C. R., & Whalen, J. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin & Review*, *8*(4), 698–707.

Darwin, C. (1871/1920). *The Descent of Man; and Selection in Relation to Sex*. (2nd ed.) New York: Appleton.

Dehaene, S., & Changeux, J. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience*, *5*(4), 390–407.

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representation of numbers in the animal and human brain. *Nature Neuroscience*, *21*(8), 355–361.

Descartes, R. (1646/1970). *Descartes' Philosophical Letters* Ed. A. Kenny. Oxford: Clarendon Press.

Descartes, R. (1649/1989). *The Passions of the Soul*. Ed. S. H. Voss. Cambridge, MA: Hackett.

Feigenson, L., & Carey, S. (2003). Tracking individuals via object-files: Evidence from infants' manual search. *Developmental Science*, *6*(5), 568–584.

Feigenson, L., Carey, S., & Hauser, M. (2002). The representations underlying infants' choice of more: Object files versus analog magnitudes. *Psychological Science*, *13*(2), 150–156.

Feigenson, L., Carey, S., & Spelke, E. (2002). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology*, *44*(1), 33–66.

Flombaum, J., & Hauser, M. under review.

Gelman, R., & Cordes, S. (2001). Counting in animals and humans. In E. Dupoux (Ed.), *Language, Brain, and Cognitive Development: Essays in Honor of Jacques Mehler* (pp. 279–302). Cambridge, MA: MIT Press.

Gallistel, C. R., & Gelman, R. (1992). Preverbal ad verbal counting and computation. *Cognition*, *44*, 43–74.

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences*, *4*(2), 59–65.

Hauser, M. D., Carey, S., & Hauser, L. B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proceedings of the Royal Society* (London), *267*, 829–833.

Hauser, M. D., MacNeilage, P., & Ware, M. (1996). Numberical representations in primates. *Proceedings of the National Academy of Science*, *93*, 1514–1517.

Jordan, K. E., & Brannon, E. M. Cardinal number representation in Rhesus monkeys. Poster presented at the annual North Carolina Cognition Group Conference, Durham, NC, February 2003.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*(2), 175–219.

Kaufman, E., Lord, M., Reese, T., & Volkman, J. (1949). The discrimination of visual number. *American Journal of Psychology*, *62*, 498–525.

Koechlin, E., Dehaene, S., & Mehler, J. (1997). Numberical transformations in five month old human infants. *Mathematical Cognition*, *3*, 89–104.

Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense: Large-number discrimination in human infants. *Psychological Science*, *14*(5), 396–401.

Locke, J. (1690/1975). *An Essay Concerning Human Understanding*, Ed. P. H. Nidditch. Oxford: Clarendon Press.

Mareschal, D., & Quinn, P. C. (2001). Categorization in infancy. *Trends in Cognitive Sciences*, *5*(10), 443–450.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*(3), 320–334.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). Multiple cues for quantification in infancy: Is number one of them? *Psychological Bulletin*, *128*(2), 278–294.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature*, *215*, 1519–1520.

Platt, J. R., & Johnson, D. M. (1971). Localization of position within a homogeneous behavior chain: Effects of error contingencies. *Learning and Motivation*, *2*, 386–414.

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, *80*(1–2), 127–158.

Rosenthal, R. (1911/1965). *Clever Hans*. Trans. Carl L. Rahn. New York: Holt, Rinehart & Winston.

Rovee-Collier, C. (1999). The development of infant memory. *Current Directions in Psychological Science*, *8*(3), 80–85.

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology*, *38*(2), 259–290.

Simon, T. J. (1997). Reconceptualizing the origins of number knowledge: A "non-numerical account." *Cognitive Development*, *12*, 349–372.

Smith, B. R., Piel, A. K., Candland, D. K. (2003). Numerity of a socially housed hamadryas baboon (Papio hamadryas) and a socially housed squirrel monkey (Saimiri sciureus). *Journal of Comparative Psychology*, *117*(2), 217–225.

Spelke, E. S. (2000). Core knowledge. *American Psychologist*, *55*(11), 1233–1243.

Starkey, P., & Cooper, R. G., Jr. (1980). Perception of numbers by human infants. *Science*, *210*, 1033–1035.

Strauss, M. S., & Curtis, L. E. (1984). Development of numerical concepts in infancy. In C. Sophian (Ed.), *Origins of Cognitive Skills* (pp. 131–156). Hillsdale, NJ: Erlbaum.

Uller, C., Carey, S., Huntley-Fenner, G., & Klatt, L. (1999). What representations might underlie infant numerical knowledge. *Cognitive Development*, *14*, 1–36.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483–488.

Whalen, J., Gelman, R., & Gallistel, C. R. (1999). Non-verbal counting in humans: The psychophysics of number representation, *Psychological Science*, *10*, 130–137.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, *358*, 749–750.

Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition*, *89*(1), B15–B25.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*, B1–B11.

7 Neural Correlates of Numerical Cognition in the Neocortex of Nonhuman Primates

Andreas Nieder and Earl K. Miller

Abstract assessment of numerical information is thought to be a phylogenetically and ontogenetically early faculty (see chapter 6); both animals (Boysen & Capaldi, 1993; Emmerton, 2001) and human infants are able to discriminate stimuli based on numerosity. This led to the hypothesis that the language-based counting and mathematical abilities found in humans build up on an evolutionarily older, nonverbal precursor system. Thus, a better understanding of the neural principles giving rise to nonverbal numerical abilities in animals will help to elucidate the neural foundation of more advanced numerical abilities only found in humans.

During the past decade, functional imaging studies (positron emission tomography, or PET; functional magnetic resonance imaging, or fMRI) provided many fruitful insights about how the human brain processes numerical information (see chapter 8). Most notably, functional imaging helped to pin down brain areas dedicated to number processing, as well as their relative contributions in different tasks (e.g., Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999; Eger, Sterzer, Russ, Giraud, & Kleinschmidt, 2003). But since such noninvasive approaches measure hemodynamic activity of relatively large brain regions, with a poor temporal resolution, it is not possible to determine the precise nature of neural activation. Thus, neurophysiology in animal models is inevitable if we want to know how single neurons in the brain give rise to numerical competence.

Rhesus monkeys (*Macaca mulatta*) are excellent model organisms to study numerosity judgments and their neural correlates. Macaques are endowed with considerable numerical competence; they can distinguish between sets of visual elements on the basis of number alone, and are even able to represent the ordinal relations between the numbers 1 to 9 (Brannon & Terrace, 1998; 2000). Elementary arithmetic abilities comparable to human babies (Wynn, 1992) have been reported for wild rhesus monkeys by Hauser and co-workers (Hauser, MacNeilage, & Ware, 1996; Hauser, Carey, & Hauser, 2000); the monkeys were able to detect simple additive and subtractive changes in the number of objects. Because of the comparatively well understood neural structures of the primate brain (see chapter 12, this volume) and the relative

similarity between the monkey and human brain (see chapters 1 and 3, this volume), macaques constitute an ideal model organism to investigate the neural substrates and mechanisms underlying numerical competence (Sawamura, Shima, & Tanji, 2002; Nieder, Freedman, & Miller, 2002; Nieder & Miller, 2003; Ninokura, Mushiake, & Tanji, 2003a,b).

Discrimination in a Delayed Match-to-Numerosity Task

Two monkeys viewed a sequence of two displays separated by a memory delay and were required to judge whether the displays contained the same small number of items (1–7) (figure 7.1A). To ensure that the monkeys solved the task by judging number per se rather than simply memorizing sequences of visual patterns or paying attention to low-level visual features that correlate with number, we employed two types of stimulus manipulations. We randomly varied the position of the items over 24 locations centered on the monkey's center of gaze as well as randomly varied the items between five different sizes. We also used eight sets of stimuli that, across them, controlled for changes in total area of the items, total circumference, density, and exact appearance (Nieder et al. 2002).

The average discrimination performance curve of both monkeys for all conditions was a smoothly declining function that is well described by a sigmoid function (figure 7.1B). Using 60 percent correct performance as criterion, the upper limit of discriminable visual quantities was between 4 and 5 items (Nieder & Miller, 2004). Thus, the animals reliably discriminated numerosities 1 to 4, but failed for numerosities of 5 and higher. However, this was only true at a numerical distance of one between match and nonmatch numerosity; if the numerical distance was increased, performance recovered (Nieder & Miller, 2003).

Prefrontal and Posterior Parietal Cortices: Candidate Structures for Numerical Processing

The lateral prefrontal (LPFC) and the posterior parietal cortices (PPC) were chosen as target areas for single unit recordings because the properties of neurons in both LPFC and PPC in monkeys suggest that these areas—which are anatomically and functionally interconnected (Quintana & Fuster, 1999; Chafee & Goldman-Rakic, 2000)—are ideal candidates to find a neural correlate for nonverbal numerosity judgments. They receive highly processed multimodal input (Miller & Cohen, 2001; see chapter 14)—a prerequisite for numerical competence because the number concept applies equally well to all sensory modalities. Both are cardinal processing stages for executive functions (e.g., working memory, decision making, goal-directed behavior, etc.) and play an important role in maintaining information "on line" (Miller & Cohen, 2001; see

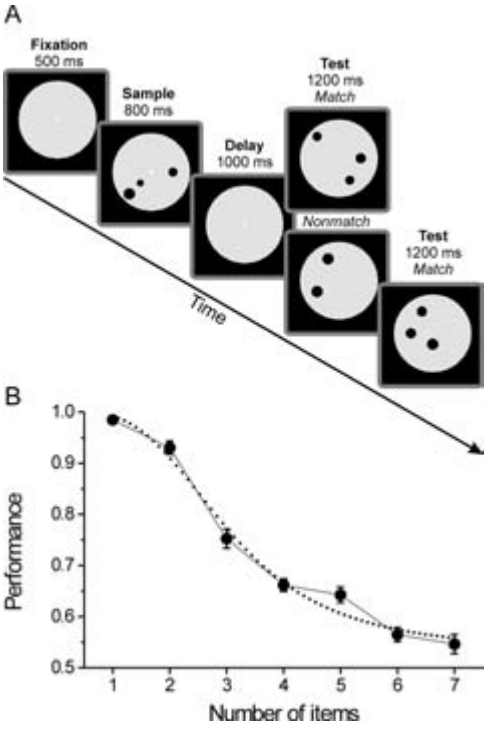


Figure 7.1

Task protocol and behavioral performance. (*A*) Stimulus protocol for the delayed match-to-sample task. A trial started when the monkey grabbed a bar. In the first 500 msec, the monkey only had to fixate a small fixation spot in the center of the display. A sample was displayed for 800 msec, followed by a 1000 msec delay period. The test stimuli contained either the same number of items (match), or one more or one less item (nonmatch) than the sample display. Matches and nonmatches appeared with equal probability. If a match appeared, the monkey had to release the lever to receive a reward. If a nonmatch appeared, the monkey had to wait for the second test stimulus (that was always a match) to get a reward for bar release. (*B*) Averaged performance of both monkeys to all stimulus conditions. The dotted line represents the best sigmoid fit to the data. Error bars: ±SEM.

chapter 15). Most interestingly, recent studies demonstrated that LPFC is involved in abstract categorization (Freedman, Riesenhuber, Poggio, & Miller, 2001), which is another prerequisite for number coding. PPC, in particular, hosts neural circuitry dedicated to the representation of abstract spatial information (chapter 14), and quantity has to be extracted from a scene containing multiple elements. Evidence for a cardinal role of LPFC and PPC is also provided by lesion and functional imaging studies in humans (see chapter 8). For example, lesions in PPC cause "acalculia," a selective deficit in arithmetic, while PFC lesions can yield specific impairment in executing numerical operations in the appropriate order. Brain imaging studies in humans revealed that primarily LPFC and PPC are activated during calculation tasks (Dehaene, 1997). Recently, it has been proposed that the human brain contains distinct neural circuits for calculation: representation of approximate quantities was found in the parietal lobe whereas the frontal lobe was predominantly active during exact calculation (Dehaene et al., 1999).

Cortical Single-Unit Recordings

Single-unit recordings were done from the same animals while they performed the delayed match-to-numerosity task. The activity of 308 LPFC and 612 PPC neurons was tested with different combinations of standard and/or control stimuli.

In LPFC, about a third (31 percent) of the randomly selected PFC cells responded selectively to the numerosity in the displays, without being affected by the stimulus type. Such abstract encoding is necessary because quantities can be represented independent of the exact appearance of the counted items. The proportion of numerosity-selective neurons has been lower in PPC, where, on average, only 9 percent encoded numerical information in a abstract way. However, visual number-encoding neurons were not uniformly distributed across the PPC (figure 7.2A). We found a clustering of numerosity neurons in the fundus of the intraparietal sulcus, which likely corresponds to the caudal part of anatomical area VIP (Nieder & Miller, submitted).

Our findings contrast a report about abundant sensorimotor number-encoding neurons in the superior parietal lobule (SPL), area 5 (Sawamura, Shima, & Tanji, 2002). Sawamura et al. (2002) trained monkeys to alternate between five arm movements of one type and five of another. They found neurons in a somatosensory-responsive region of the SPL that kept track of the movement number. Relatively few such neurons were found in the same lateral PFC regions where other perceptual categories have been found. One possibility for the difference between these studies may be modality (touch vs. vision), but another may be the level of abstraction. Most movement-number representations found by Sawamura and colleagues (85 percent) were not abstract; number-selective activity depended on whether the monkey's

movement was "push" or "turn." By contrast, the visual number representations found in the LPFC were abstract and generalized. Changes in the physical appearance of the displays had little effect on activity of the majority of number-tuned neurons (figure 7.2B–E, plate 14).

Apart from the proportions of numerosity-encoding neurons, which were significantly greater in the LPFC than in any PPC region, the basic tuning properties of LPFC and PPC neurons were comparable and will, thus, be treated together. Neurons in both areas showed a significant decrease in the activity on error trials, both in the sample and delay interval, which suggests their direct involvement in task performance. Numerosity-selective neurons were "tuned" for the number of items on a visual display, i.e., they showed maximum activity to one of the five presented quantities (a neuron's "preferred numerosity"). Across the population, "one" was the number most often preferred; neural preference was distributed equally among the remaining numbers.

Numbers are not isolated categories, but exist in relation to one another (e.g., "3" is greater than "2" and less than "4"). Thus, a defining characteristic of numerical competence is ordinality (Gelman & Gallistel, 1978). Neural activity in the PFC seemed to preserve numerical order; neurons showed peak activity to a specific number and a progressive drop off as number progressively varied (figures 7.2B–E, 7.3C, D). On average, activity dropped off progressively with number for both the sample and delay intervals. The normalized and averaged neural activity of all neurons in both LPFC and PPC formed a bank of overlapping numerosity filters (figure 7.3C). The properties of these filter functions can explain fundamental phenomena that both humans and animals share when extracting numerical information.

Numerical Distance Effect

Discrimination between two numbers improves with increasing numerical distance between them. For example, it is easier to discriminate 2 and 6 than it is 5 and 6. This effect is known as the *numerical distance effect* and has been found both in animals and humans when judging numerosities (Mechener, 1958; Moyer & Landauer, 1967; Van Oeffelen & Vos, 1982; Dehaene, Dehaene-Lambertz, & Cohen, 1998; Brannon & Terrace, 2000).

In a different set of experiments (Nieder & Miller, 2003), we therefore tested the monkeys' performance to an expanded range of nonmatch numerosities (figure 7.3A, B). Monkeys made more errors when the nonmatch numerosities were adjacent to the sample numerositiy, and they performed progressively better as numerical distance between two displays increased. This numerical distance effect found in the monkeys' behavior can be explained by the neural filter properties considering simple signal detection principles.

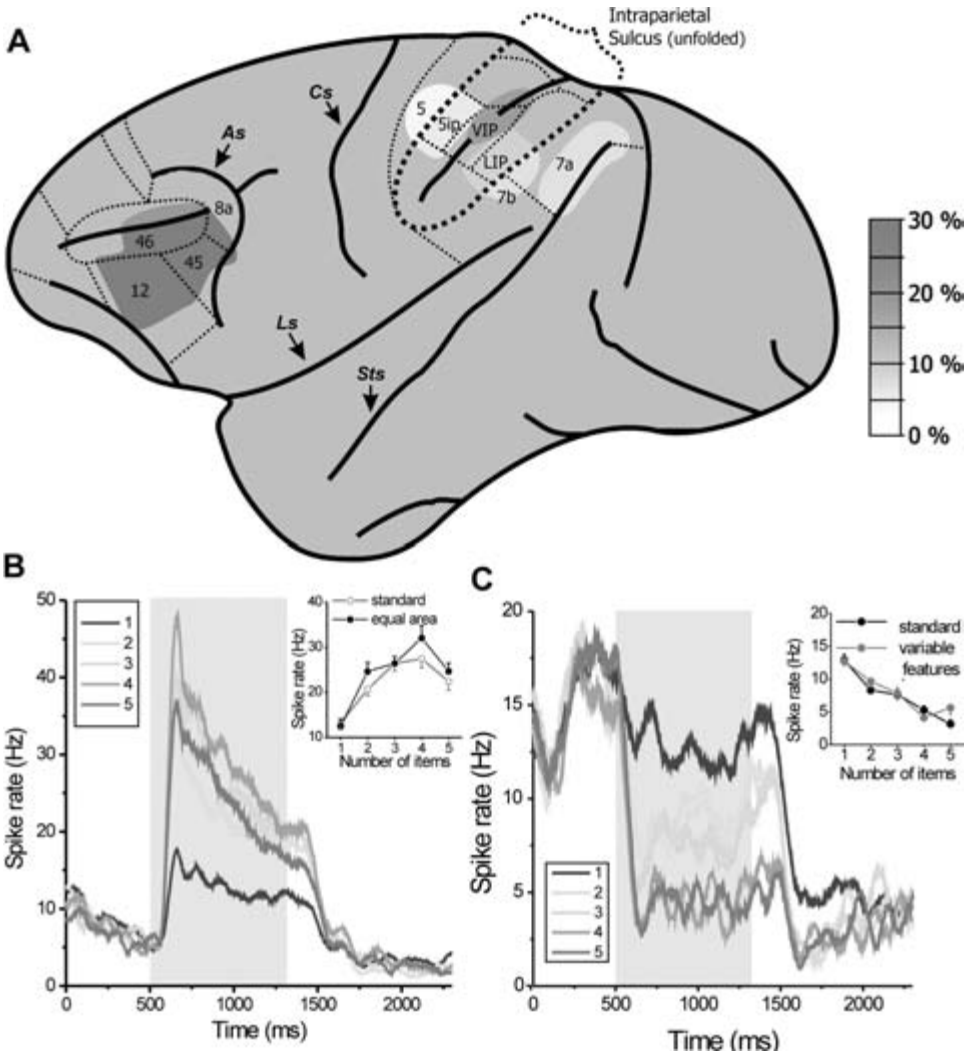


Figure 7.2

Recording sites and neural responses. (*A*) Recordings sites in the lateral prefrontal and posterior parietal cortices (lateral view of a monkey brain). The relative proportions of numerosity selective neurons in each area are color coded. Color legend to the right of the brain. 5, area 5; 5ip, intraparietal part of area 5; VIP, ventral intraparietal area; LIP, lateral intraparietal area; 7a, area 7a; 7b, area 7b; Cs, central sulcus; Ls, lunate sulcus; Sts, superior temporal sulcus (Nieder & Miller, submitted). (*B–E*) Spike density histograms showing the responses of four example neurons from the LPFC (*B*, *D*) and intraparietal sulcus (*C*, *E*) to numerosities. Neurons were selective during the sample period (*B*, *C*) and/or the delay interval (*D*, *E*). Each colored line shows the time course of activity for the five tested numerosities (standard and control stimuli pooled). The insets indicate the mean spike rate of the neurons to the standard and the control stimulus. Error bars: SEM. See plate 14 for color version.

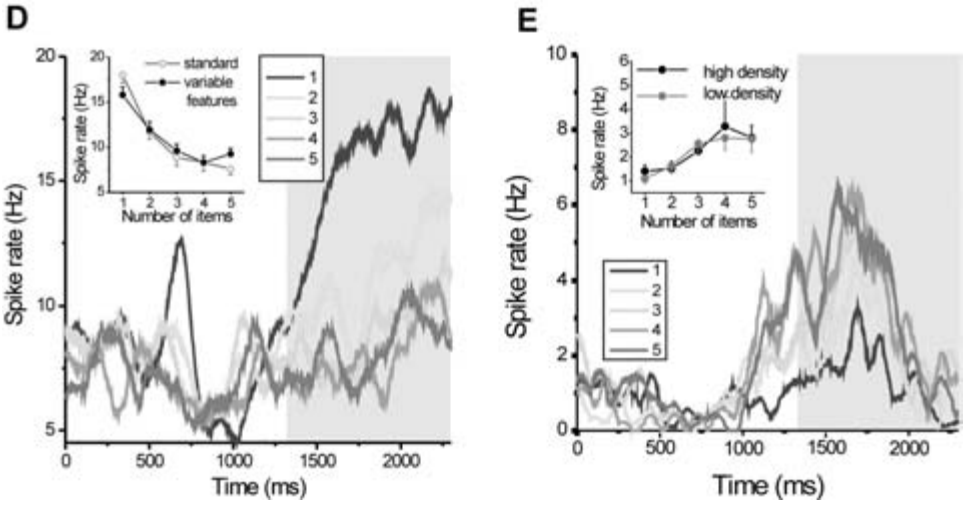
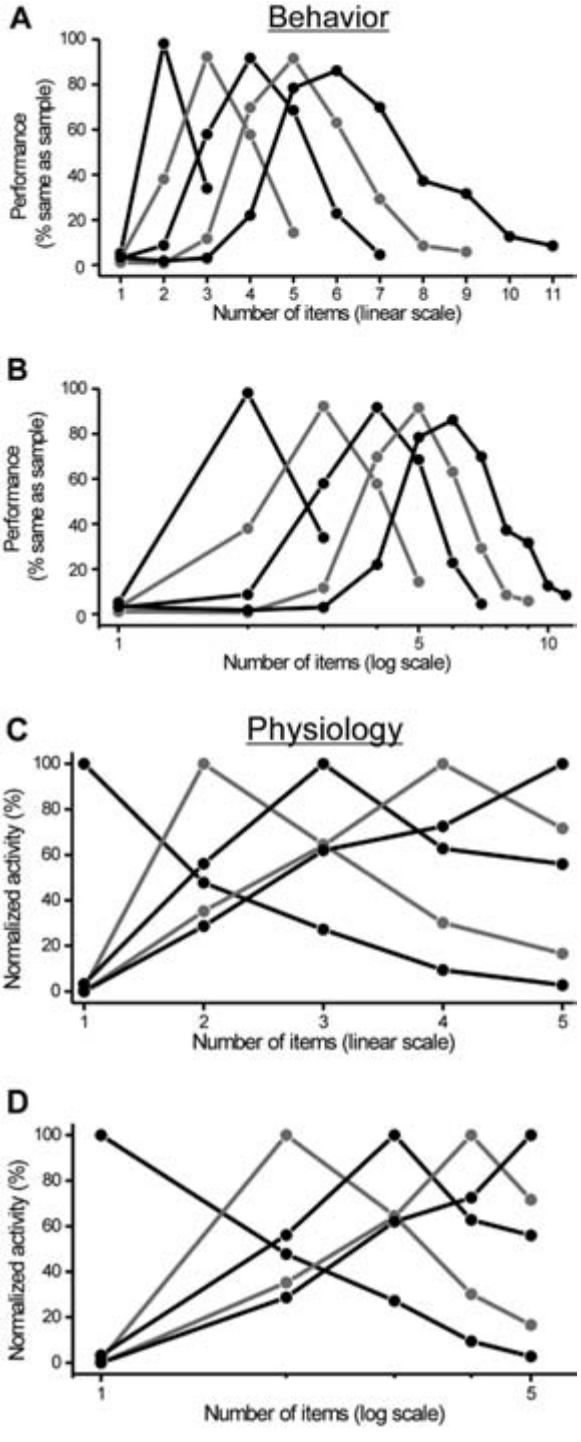


Figure 7.2
(continued)

Within the framework of signal detection theory (Green & Swets, 1966), the decisions about whether a stimulus is numerosity 1 or 4, for example, is based on whether this stimulus produces a greater response in a neuron selective for numerosity 1 or 4, respectively. Because the tuning curves of neurons selective for 1 or 4 are barely overlapping (figure 7.3C, D), numerosity 4 will almost always generate a larger response in neurons tuned to numerosity 4, leading to an almost perfect discrimination performance for numerosity 1 versus 4. However, because the tuning functions of neurons selective for, say, 3 and 4 are strongly overlapping, numerosity 3 will be more likely to evoke a response in the neurons selective for numerosity 4. Thus, on single trials, numerosity 3 may evoke a greater spike count in filters tuned to "4" than numerosity 4 does, and as a consequence the monkeys will make many errors when discriminating numerosity 3 and 4, or any other adjacent numerosity pairs.

Numerical Magnitude Effect

Another very basic behavioral effect can be explained as well by the neural tuning properties. The discrimination of two numbers of a given numerical distance gets more difficult the higher the absolute values of the two numbers. For example, it is easier to discriminate 2 and 3, than 5 and 6, even though the numerical distance is 1 in both cases. This *numerical magnitude effect* was also present in the monkeys' performance. For larger quantities, the two numerosities had to be numerically more distant for performance to reach the level obtained with smaller quantities and closer numerical distance (figure 7.3A, B). This behavioral effect can also be explained by the


A
Behavior
Performance (% same as sample)
Number of items (linear scale)
B
Performance (% same as sample)
Number of items (log scale)
C
Physiology
Normalized activity (%)
Number of items (linear scale)
D
Normalized activity (%)
Number of items (log scale)

neuronal tuning curves (Nieder et al., 2002). The average bandwidth of the neural filters increased with quantity (i.e., on average, neurons became less precisely tuned as their preferred quantity increased) (figure 7.3C, D). Thus, if a monkey has to discriminate small numerosities (say, 1 and 2), rather selective neural filters become engaged that do not overlap a lot. Thus the signal-to-noise ratio of a discrimination of 1 versus 2 will be high, resulting in few errors. On the other hand, if a monkey wants to discriminate large numerosities (like 4 and 5), the filter functions will overlap considerably. Therefore, a discrimination of 4 versus 5 will show a low signal-to-noise ratio, leading to a rather poor performance.

Weber's Law

Numerical judgments are clearly different from sensory processes because they can be abstract, irrespective of exact physical appearance. At the same time, the numerical distance and magnitude effects illustrate that numerical judgments exhibit phenomena typically found for sensory judgments. These similarities are not superficial, but include fundamental psychophysical laws. One such example is Weber's law of "just noticeable differences" (Weber, 1850). The difference threshold (or "just noticeable difference," ΔI) is the minimum amount by which stimulus magnitude must be changed in order to produce a noticeable variation in sensory experience. It scales with stimulus magnitude and thus, the so-called Weber fraction ($\Delta I/I$) is a constant.

If Weber's law should hold for numerical judgments, then it is important to demonstrate that the minimum numerical distance needed by the monkeys to reach a certain discrimination performance grows proportionally with the reference numerosity (or sample numerosity, respectively). That way, the *just noticeable numerical difference* (ΔI) divided by the sample numerosity (I) should stay constant across all tested numerosities. Even more, the same should hold true for the neural numerical representations, the numerosity tuning curves.

◀ **Figure 7.3**

Behavioral and neural numerical filter function in the LPFC. (*A*, *B*) The behavioral performance for both monkeys indicated whether they judged the first test stimulus (after the delay) as containing the same number of items as the sample display ("% same as sample"). The function peaks indicate the sample numerosity at which each curve was derived. Behavioral filter functions are plotted on a linear (*A*) and logarithmic (*B*) scale. The functions are asymmetric when plotted on a linear scale (note the shallower slope towards higher numerosities) (*A*), but are symmetric when plotted on the nonlinear logarithmic scale (*B*). (*C*, *D*) Neural representation of numerosities during the sample period. (*C*) The neural filter functions are asymmetric on a linear scale (note the shallower slope toward higher values for preferred numerosity "2," for example). (*D*) Logarithmic transformation of the filter functions results in more symmetric distributions. Error bars: ±SEM.

We calculated the Weber fraction according to the equation by Van Oeffelen & Vos, (1982) (see Nieder & Miller, 2003). Indeed, with a mean value of 0.35, the Weber fractions derived from the behavioral filter functions (figure 7.3A, B) were constant across numerosities. More importantly, even the neural filter functions exhibited constant Weber fractions within the range of tested numerosities. When comparing behavioral and neural filter functions, a constant bandwidth ratio of 1.5 across all numerosities was observed. This indicates a direct relationship between behavioral and neuronal representations, with a greater sensitivity (by a factor of 1.5) on the behavioral than the neural level.

Scaling of the "Mental Number Line"

The similarities between sensory and cognitive magnitudes are even more striking. Judgments of the magnitude of sensory stimuli are not linearly scaled. Rather, perceptual representations are nonlinearly compressed. Fechner (1860) suggested a logarithmically compressed scale, so that linear increments in sensation S are proportional to the logarithm of stimulus magnitude I (Fechner's law, $S = k \cdot log(I)$). Stevens (1961) instead postulated that sensation is a power function of the stimulus magnitude (Stevens' law, $S = k \cdot I^n$). Both Fechner's law and Stevens' law are largely valid for general sensory phenomena and also account for many properties of sensory neurons (Dayan & Abbott, 2001). Therefore, if a continuum between perceptual and cognitive processes exists, numerical representations should also be scaled on a nonlinear, compressed "number line."

Signal detection tasks, like judging numerosity, yield response probability density functions that are normal distributions (Gauss functions). Behavioral discrimination and single-unit tuning functions (figure 7.3) can be regarded as the monkeys' behavioral and neural probability density functions, or numerical representations, respectively. The important question now is which scaling scheme provides normal, i.e., symmetric distributions of the numerical representations. To that aim, we (Nieder & Miller, 2003) plotted the behavioral and neural filter functions along different scales and analyzed which scaling scheme gave the most symmetric distributions. It turned out that both the performance and the single unit data for numerosity judgments are better described using a power function–compressed (Stevens' law) or logarithmically compressed (Fechner's law) scale, as opposed to a linear scale. Thus, abstract numerical representations in monkeys obey basic psychophysical laws that postulate a nonlinearly compressed scaling of sensory experience. In other words, the nonverbal "number line" is not linear but non-linearly compressed. Interestingly, nonlinearly compressed coding of numerical representations has also been postulated based on neural network simulations (Dehaene & Changeux, 1993; Verguts & Fias, 2004).

Why would such a compressed "number line" be useful? One advantage of a compressed scaling would be that higher numbers could be represented with a smaller

proportion of neurons. The trade-off of this mechanism is that discrimination of high numbers is more difficult and less precise than the discrimination of small numbers, a phenomenon we experience in everyday life when we have to rapidly judge a number of objects in situations where we lack time for counting.

Because sensory and cognitive magnitude judgments obey fundamental psychophysical laws, we thus suggest that perhaps cognition is built upon principles that were originally developed for lower-level abilities, like perception. If this is the case, then higher-level cognitive tasks found only in humans, like complex counting, may share some fundamental properties with more basic processes that are present in many kinds of creatures.

A Parietofrontal Number Network

If both LPFC and PPC harbor numerosity selective cells, what may be the relative contributions of these areas? Is there a functional cortical hierarchy in the processing of numerical information? To address such questions, we analyzed the neural response properties in more detail and directly compared LPFC and PPC activity.

We found a difference in the neural latencies of number-encoding LPFC and PPC neurons. With a median of 88 ms, response latencies were significantly faster in the PPC compared to the LPFC with a median of 120 ms (figure 7.4A, B). More importantly, PPC neurons not only activated faster, they also began to exhibit numerosity selectivity prior to PFC (figure 7.4C, D). A sliding statistical comparison (Kruskal-Wallis test) revealed that PPC neurons (median selectivity latency: 99 ms), on average, discriminated between numerosities 17 ms earlier than PFC neurons (116 ms; $p < 0.05$, Mann Whitney U test, two-tailed). The tuning strengths of numerosity selective neurons (defined by $(R_{max} - R_{min})/(R_{max} + R_{min})$, where R_{max} and R_{min} are the maximum and minimum mean spike rates) were equal in the LPFC and PPC during viewing (sample epoch average tuning index: 0.33 and 0.31 in PPC and LPFC, respectively), but significantly stronger in the LPFC during the memory delay (PPC average index = 0.25, LPFC average index = 0.32; $p < 0.05$, Mann Whitney U test, two-tailed) (figure 7.4E).

These data help bring the roles of the PPC and LPFC into clearer focus. The shorter PPC response and selectivity latencies suggest that it may be the first cortical stage that extracts visual numerical information. As PPC and LPFC are functionally interconnected, that information may be conveyed to the LPFC where it is amplified and maintained (hence the greater incidence of effects and stronger tuning in the LPFC during the memory delay) to gain control over behavior. Whether there is clear-cut modality-specific numerical processing in the PPC (VIP for visual-number, the SPL for sensorimotor-number) needs to be determined in future multimodal approaches. Area VIP, by all means, would be an ideal candidate structure to integrate supramodal

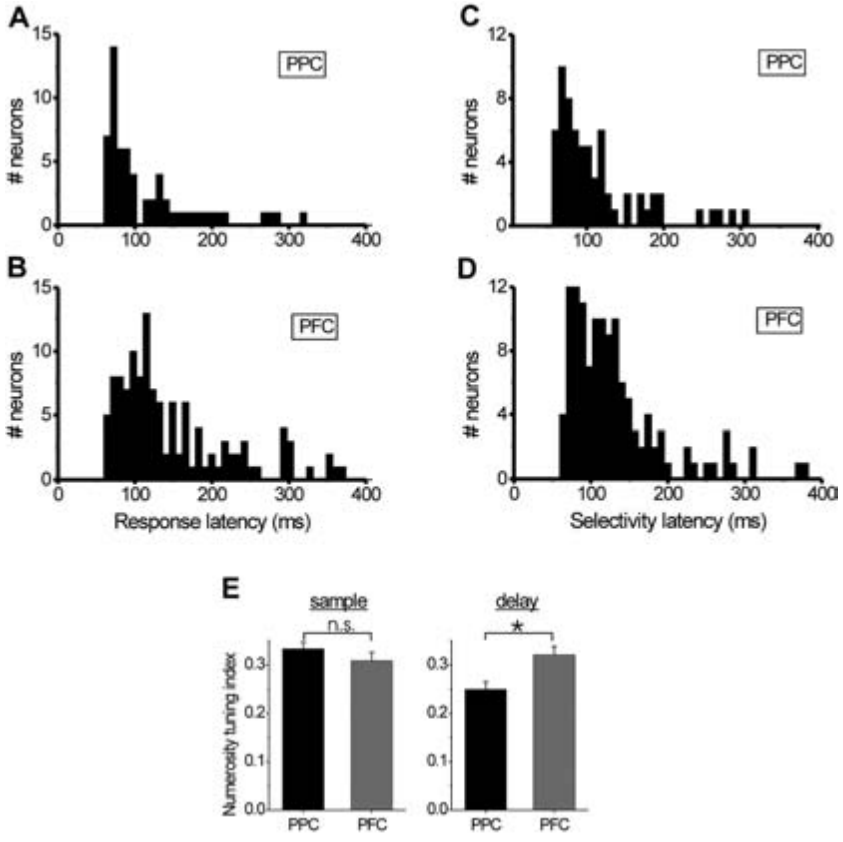


Figure 7.4

Comparison of numerosity selective neurons in the prefrontal and posterior parietal cortices. (*A*, *B*) Response latency histogram of neurons in the posterior parietal and prefrontal cortices. (*C*, *D*) Distributions of numerosity selectivity latencies in the posterior parietal and prefrontal cortices. (*E*) Numerosity tuning indices for prefrontal and posterior parietal neurons compared in the sample and delay interval.

numerical information, since it integrates visual, auditory, and somatic input. Also, whether the lack of sensorimotor number-tuning in the LPFC may reflect modality-specificity or other factors like the aforementioned difference in level of abstraction remains to be determined. The latter possibility is suggested by the greater abstraction of perceptual categories in the PFC than in the temporal cortical areas that provide it with visual input (Freedman, Riesenhuber, Poggio, & Miller, 2003). Finally, our data indicate close homologies between humans and monkeys. Imaging studies suggest prime involvement of LPFC and PPC (particularly the ventral intraparietal sulcus) for "number sense" in humans (chapter 8). The clustering of neurons in a corresponding region in monkeys and their shorter response and selectivity latencies than LPFC neurons is consistent with it being a prime source of number information in both humans and nonhuman primates.

References

Boysen, S. T., & Capaldi, E. J. (Eds.) (1993). *The Development of Numerical Competence: Animal and Human Models*. *Hillsdale*, NJ: Erlbaum.

Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science*, *282*, 746–749.

Brannon, E. M., & Terrace, H. S. (2000). Representation of the numerosities 1–9 by rhesus macaques (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*, *26*, 31–49.

Chafee, M. V., & Goldman-Rakic, P. S. (2000). Inactivation of parietal and prefrontal cortex reveals interdependence of neural activity during memory-guided saccades. *Journal of Neurophysiology*, *83*, 1550–1566.

Dayan, P., & Abbott, L. F. (2001). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Cambridge, MA: MIT Press.

Dehaene, S. (1997). *The Number Sense: How the Mind Creates Mathematics*. Oxford: Oxford University Press.

Dehaene, S., & Changeux, J. P. (1993). Development of elementary numerical abilities: A neural model. *Journal of Cognitive Neuroscience*, *5*, 390–407.

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representation of numbers in the animal and human brain. *Trends in Neurosciences*, *21*, 355–361.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: behavioral and brain-imaging evidence. *Science*, *284*, 970–974.

Eger, E., Sterzer, P., Russ, M. O., Giraud, A. L., & Kleinschmidt, A. (2003). A supramodal number representation in human intraparietal cortex. *Neuron*, *37*, 719–725.

Emmerton, J. (2001). Avian visual cognition. Online. Available: www.pigeon.psy.tufts.edu/avc/emmerton/.

Fechner, G. T. (1860). *Elemente der Psychophysik*, vol. 2. Leipzig: Breitkopf & Härtel.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science*, *291*, 312–316.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2003). A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *Journal of Neuroscience*, *23*, 5235–5246.

Gelman, R., & Gallistel, C. R. (1978). The child's understanding of number. Cambridge, MA: Harvard University Press.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.

Hauser, M. D., Carey, S., & Hauser, L. B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proceedings of the Royal Society of London: Biological Sciences*, *267*, 829–833.

Hauser, M. D., MacNeilage, P., & Ware, M. (1996). Numerical representations in primates. *Proceedings of the National Academy of Sciences*, *93*, 1514–1517.

Mechner, F. (1958). Probability relations within response sequences under ratio reinforcement. *Journal of the Experimental Analysis of Behaviour*, *1*, 109–121.

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

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature*, *215*, 1519–1520.

Nieder, A., Freedman, D. J., & Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science*, *297*, 1708–1711.

Nieder, A., & Miller, E. K. (2003). Coding of cognitive magnitude: Compressed scaling of numerical information in the primate prefrontal cortex. *Neuron*, *34*, 149–157.

Nieder, A., & Miller, E. K. (2004). Analog numerical representations in rhesus monkeys: Evidence for parallel processing. *Journal of Cognitive Neuroscience* (in press).

Ninokura, Y, Mushiake, H., & Tanji, J. (2003a). Representation of the temporal order of visual objects in the primate lateral prefrontal cortex. *Journal of Neurophysiology*, *89*, 2868–2873.

Ninokura, Y., Mushiake, H., & Tanji, J. (2003b). Integration of temporal order and object information in the monkey lateral prefrontal cortex. *Journal of Neurophysiology* (in press).

Quintana, J., & Fuster, J. M. (1999). From perception to action: Temporal integrative functions of prefrontal and parietal neurons. *Cerebral Cortex*, *9*, 213–221.

Sawamura, H., Shima, K., & Tanji, J. (2002). Numerical representation for action in the parietal cortex of the monkey. *Nature*, *415*, 918–922.

Stevens, S. S. (1961). To honor Fechner and repeal his law. *Science*, *133*, 80–86.

Van Oeffelen, M. P., & Vos, P. G. (1982). A probabilistic model for the discrimination of visual number. *Perception & Psychophysics*, *32*, 163–170.

Verguts, T., & Fias, W. (2004). Representation of number in animals and humans: A neural model. *Journal of Cognitive Neuroscience* (in press).

Weber, E. H. (1850). Der Tastsinn und das Gemeingefühl. In R. Wagner (Ed.), *Handwörterbuch der Physiologie*, (vol. 3, part 2, pp. 481–588).

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, *358*, 749–750.

8 Evolution of Human Cortical Circuits for Reading and Arithmetic: The "Neuronal Recycling" Hypothesis

Stanislas Dehaene

Humans have a remarkable ability to invent symbol systems such as Arabic numerals or the alphabet. This capacity is unique in the animal kingdom. Thus, one has to ask what is so special about the human brain that allows it to expand its functionality by acquiring new cultural tools.

A first possibility is that, relative to other animals, the human brain has evolved new specialized processors, each providing access to a new cognitive function. For instance, our species may have evolved some special brain mechanisms for recursion that would give us access to the domain of syntax (Hauser, Chomsky, & Fitch, 2002). However, such a possibility is excluded for recent cultural acquisitions such as reading or arithmetic. Those activities are far too recent to have exerted any evolutionary pressure on brain evolution. Reading, for instance, was invented only 5,400 years ago, and symbolic arithmetic is even more recent: the Arabic notation and most of its associated algorithms were not available even a thousand years ago. Thus, it is logically impossible that there exist dedicated brain mechanisms evolved for reading or symbolic arithmetic.

An alternative theory is that those capacities rely upon an extended range of cortical plasticity unique to humans. According to this second hypothesis, the human brain would be special in its capacity to accommodate a broad range of new functions through learning. At one extreme, it may be suggested that the architecture of our brain exerts little or no constraints on the range of competences that we can acquire, because we are equipped with broad if not universal mechanisms of learning. Although admittedly presented here in somewhat caricatured form, this view is not so distant from some modern connectionist or neoconstructivist statements (e.g., Quartz & Sejnowski, 1997). Such a learning-based theory might explain the vast range of human cultural abilities, but it implies that the brain implementation of those abilities should be highly variable across individuals. Depending on an individual's learning history, the same brain region might become involved in various functions. During learning, random symmetry breaking might ultimately lead to the

assignment of dedicated territories to different competences, but this assignment should be randomly determined for different individuals. Thus, one would not expect to find reproducible cerebral substrates for recent cultural activities such as reading and arithmetic.

The purpose of the present chapter is to examine where the data stand. A wealth of recent neuroimaging and neuropsychological findings shed light on the ability of the human brain to acquire novel cultural objects such as reading and arithmetic. As we shall see, those data go against the hypothesis of an unbiased, random symmetry-breaking theory of cultural learning. Converging psychological, neuropsychological and brain-imaging evidence demonstrates that the adult human brain houses dedicated mechanisms for reading and arithmetic. Small cortical regions, which occupy reproducible locations in different individuals, are recruited by these tasks. They accomplish their function automatically and often without awareness. Furthermore, the lesion of those regions can lead to specific reading or calculation impairments. In brief, the evidence seems to support the existence of distinct, reproducible and rather specific brain bases for reading and arithmetic.

The paradox, of course, is that given the available evolutionary time, it is impossible that the architecture of our brains has somehow adapted to the specific problems posed by these cultural tools. Closer examination of the function of the relevant brain areas, however, suggests a possible resolution of this paradox. It is not the case that those areas acquire an entirely distinct, culturally arbitrary new function. Rather, they appear to possess, in other primates, a prior function closely related to the one that they will eventually have in humans. Furthermore, many of the functional features that make them highly efficient in processing human cultural tools are already present. Thus, relatively small changes may suffice to adapt them to their new cultural domain.

I conclude the chapter by tentatively proposing the "neuronal recycling" hypothesis: the human capacity for cultural learning relies on a process of preempting or recycling preexisting brain circuitry. According to this third view, the architecture of the human brain is limited and shares many traits with other nonhuman primates. It is laid down under tight genetic constraints, yet with a fringe of variability. I postulate that cultural acquisitions are only possible insofar as they fit within this fringe, by reconverting preexisting cerebral predispositions for another use. Accordingly, cultural plasticity is not unlimited, and all cultural inventions should be based on the preemption of preexisting evolutionary adaptations of the human brain. It thus becomes important to consider what may be the evolutionary precursors of reading and arithmetic.

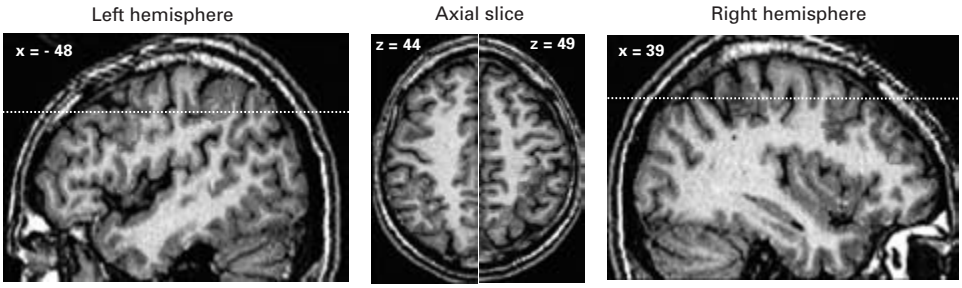
Cerebral Bases of Arithmetic

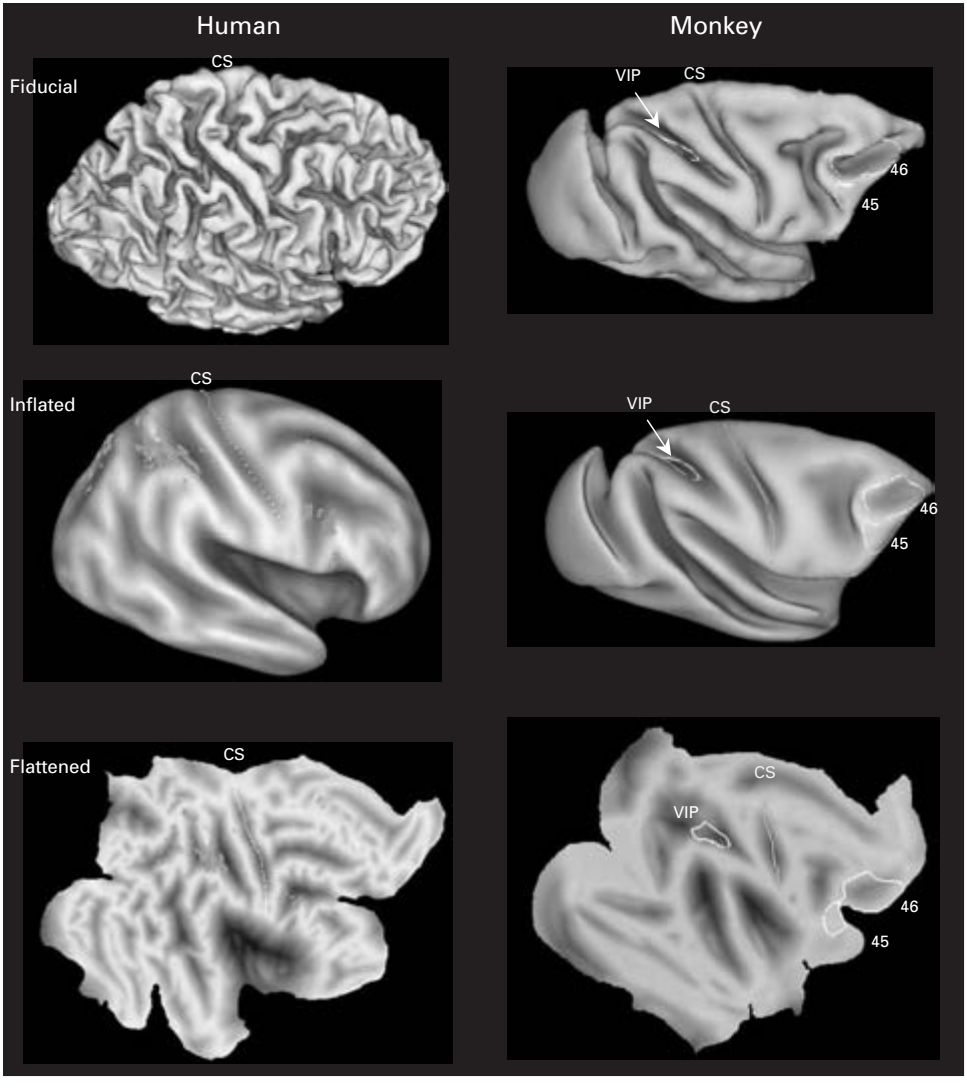
Calculation and the Human Intraparietal Sulcus

Convergent imaging and neuropsychological results associate mental arithmetic with the parietal lobe (Dehaene, Piazza, Pinel, & Cohen, 2003). The left and right intraparietal regions are systematically activated whenever subjects engage in calculation (Chochon, Cohen, van de Moortele, & Dehaene, 1999; Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999; Fias, Lammertyn, Reynvoet, Dupont, & Orban, 2003; Gruber, Indefrey, Steinmetz, & Kleinschmidt, 2001; Lee, 2000; Pesenti, Thioux, Seron, & De Volder, 2000; Rickard et al., 2000; Simon, Mangin, Cohen, Le Bihan, & Dehaene, 2002; Zago et al., 2001). Their degree of activation is directly proportional to the difficulty of the arithmetic task, as measured by the size of the numbers involved, the numerical distance that separates them, or the number of operations to be performed in a given time (Menon, Rivera, White, Glover, & Reiss, 2000; Pinel, Dehaene, Riviere, & LeBihan, 2001; Stanescu-Cosson et al., 2000).

The interpretations of these findings, however, remains debated. According to one theory, a bilateral subregion of the parietal lobe, located deep inside the intraparietal sulcus, contains a domain-specific representation of numerical quantity (Dehaene & Cohen, 1995; Dehaene et al., 2003). An alternative "domain-general" view, however, proposes that no specific representation is dedicated to number processing (Simon, 1999), and that the engagement of the parietal lobe during calculation can be explained entirely by concomitant task components such as finger counting and visuospatial working memory (Gruber, Indefrey, Steinmetz, & Kleinschmidt, 2001; Zago et al., 2001).

Several findings support the domain-specific hypothesis. First, when multiple visuospatial, language, and calculation tasks are imaged in the same subjects, a small subregion in the depth of the horizontal segment of the intraparietal sulcus (hereafter called the HIPS region), is found active solely during calculation (Simon et al., 2002). Thus, its activation cannot be reduced to spatial, attentional, eye or finger movement artifacts. Second, the HIPS activates when subjects merely have to detect Arabic numerals, but not letters or colors, in a stream of auditory or visual stimuli (Eger, Sterzer, Russ, Giraud, & Kleinschmidt, 2003). This indicates that neither calculation nor working memory are needed to obtain parietal number-related activations. Indeed, the HIPS is even activated by subliminal numerals, indicating automatic access to quantity information from number symbols (Naccache & Dehaene, 2001). Third, a recent meta-analysis (Dehaene, et al., 2003) indicates that the HIPS is jointly activated by esentially all number processing contrasts that have been used in the literature, and particularly when the task puts emphasis on quantity processing (figure 8.1, plate 15). It is unlikely that all experiments are affected by identical artifacts, especially considering that some studies have contrasted highly similar tasks with the same difficulty


Left hemisphere
Axial slice
Right hemisphere
x = - 48
z = 44
z = 49
x = 39


Human
Monkey
Fiducial
Inflated
Flattened
CS
VIP
46
45

level as measured by response time and error rate (e.g., approximation relative to exact calculation: Dehaene et al., 1999).

Further evidence for the tight relation between the HIPS and mental arithmetic comes from the classical neuropsychological finding that lesions to the left parietal cortex cause severe impairments in calculation, sometimes without much concomitant cognitive impairment in other domains of reasoning (e.g., Dehaene & Cohen, 1997; Lee, 2000; Takayama, Sugishita, Akiguchi, & Kimura, 1994). Brain imaging and neuropsychological evidence points to even more selective dissociations, for instance, between subterritories for subtraction and multiplication (Duffau et al., 2002; Lee, 2000).

Precursors of Arithmetic in Animals

In the last decades, the systematic investigation of precursors of numerical abilities in animals has shed some light on the biological origins of human arithmetic. Behavioral investigations have revealed that animals such as rats, pigeons, or monkeys can extract the approximate numerosity of auditory or visual sets of objects (see, e.g., chapters 6 and 7). Numerosity is represented by animals independently of other parameters such as object size or shape (Brannon & Terrace, 1998). Evidence from wild animals indicates that numerosity is part of the spontaneous representational repertoire of many animal species, and does not need to be inculcated by training (Hauser, Carey, & Hauser, 2000; Hauser, Dehaene, Dehaene-Lambertz, & Patalano, 2002; McComb, Packer, & Pusey, 1994). Even when training is involved in laboratory animals, experiments have demonstrated generalization patterns that goes beyond what the animals could have acquired by mere stimulus-driven learning. For instance, macaques trained to order the numerosities 1 though 4 generalized spontaneously to the range of numbers 5 through 10 (Brannon & Terrace, 1998). Likewise, macaques trained on a

◄ **Figure 8.1**
Core regions for number processing in humans. The image on top shows the intersection of activations observed in several tasks including number comparison, simple arithmetic, approximate calculation, and subliminal quantity processing (adapter from a meta-analysis in Dehaene et al., 2003). Activations are systematically observed in the bilateral horizontal segment of the intraparietal sulcus (HIPS) as well as in precentral cortex. Caret software (Van Essen et al., 2001) was used to map the observed activations onto an unfolded map of the human cortex (only the right hemisphere is shown). For comparison, similar views of the macaque brain are shown, with white borders indicating the areas where neurons tuned to numerosity have been found (see chapter 7, this volume). The human HIPS region, in the depth of the intraparietal sulcus, is a plausible homolog of the macaque area VIP. The human precentral activation is more distant from the monkey areas 45/46, suggesting a greater amount of distortion in prefrontal cortex during evolution, as also indicated by other comparative studies (Courtney, Petit, Maisog, Ungerleider, & Haxby, 1998; Nakahara, Hayashi, Konishi, & Miyashita, 2002). See plate 15 for color version.

matching-to-sample task with training stimuli where number and total size were confounded later generalized on the basis of number, not total size when the two parameters were unconfounded (Nieder, Freedman, & Miller, 2002). Finally, there is some evidence that animals can use these number representations for simple approximate calculations such as addition or subtraction (e.g., Hauser et al., 2000).

Crucial to the link between animal research and human neuroimaging studies of arithmetic is the recent finding of a neurobiological substrate for animal number processing. In agreement with the predictions of a neural network model (Dehaene & Changeux, 1993), neurons tuned to numerosity were recently recorded in macaque monkeys trained to perform numerosity-dependent motor or matching tasks (Nieder et al., 2002; Nieder & Miller, 2003; chapter 7, this volume; Sawamura, Shima, & Tanji, 2002). One such neuron might respond to visual displays of three objects, regardless of their spatial organization, size or shape, while responding much less to two or four objects and not at all to one object or to five objects. Crucially, such neurons are found in dorsolateral prefrontal cortex, but also in the vicinity of the intraparietal sulcus, with the latest evidence suggesting a precise localization in the depth of the middle portion of the IPS, possibly within area VIP (see chapter 7). As demonstrated in figure 8.1, this localization in macaques constitutes a plausible homolog of the human site of activation during symbolic arithmetic tasks. It was indeed predicted on the basis of the localization of human arithmetic-related activations relative to putative human homologs of parietal areas LIP and AIP (Simon et al., 2002).

A Similar Principle of Numerosity Tuning in Monkeys and Humans

In order to demonstrate that the monkey competence for approximate numerosity representation is a plausible precursor of human arithmetic, one should ideally show that the human HIPS region also contains numerosity-sensitive neurons. Yet most human neuroimaging studies have used symbolic stimuli (e.g., Arabic digits) and have not probed basic numerosity processing. Furthermore, single neurons are not accessible in humans with noninvasive methods. Recently, however, my colleagues and I have used an habituation design to demonstrate numerosity tuning in the human HIPS (Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004).

We recorded whole-brain fMRI images continuously while subjects were repeatedly presented with many visual displays of a fixed, large numerosity (e.g., always 16 dots). We reasoned that this should lead to habituation of a coarse population of numerosity detectors tuned around the numerosity 16. We then "read out" this state of habituation by recording the event-related fMRI activation to a single deviant numerosity (ranging from 8 to 32 dots). As predicted, the only regions that responded to numerosity change were the left and right intraparietal sulci, including the right HIPS. Activation in those regions followed an inverse Gaussian function centered around the

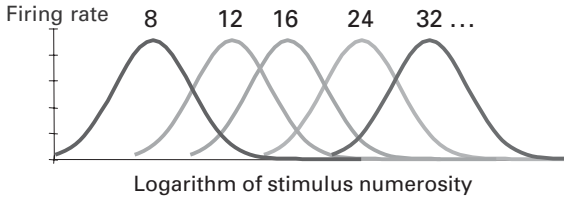
habituated numerosity (figure 8.2): it was low for deviant numerosities that fell close to the habituation numerosity (e.g., 16 followed by 20), but increased on either side as the deviant numerosity became more distant (e.g., 16 followed by 8 or by 32). This experimental design allowed us to indirectly trace the average tuning curve of the underlying neural population. The details of this fMRI tuning function were identical to those observed in the monkey: (1) Tuning was independent of the presence or absence of a concomitant change in object shape; (2) Tuning width doubled when the habituation numerosity doubled, indicating that the representation of numerosity follows Weber's law: the precision of the representation decreases linearly with the size of the numbers involved; (3) As in single-cell data, the measured responses were best described by a Gaussian tuning curve with a fixed width on a logarithmic number line.

Learning to Map Arabic and Verbal Symbols onto the Quantity Code

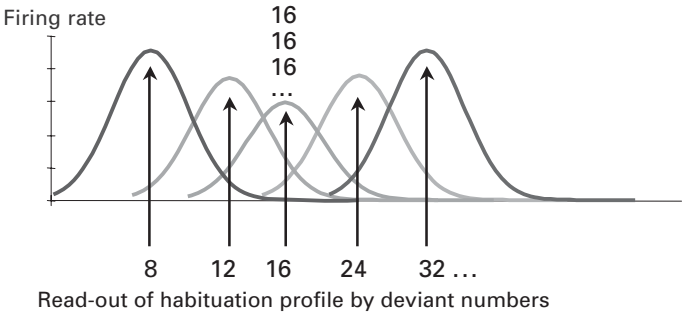
The presence of an evolutionary precursor of arithmetic in animals helps resolve the paradox described in the introduction, by demonstrating that arithmetic is not, after all, a completely arbitrary cultural invention. Although the particular symbols and algorithms that we use are conventional, our very ability to invent them rests on an intuitive understanding or "core knowledge" of the number domain, which has been termed "number sense" (Dehaene, 1997). My hypothesis is that the human acquisition of Arabic numerals and arithmetic is possible, and occurs with a reproducible underlying brain substrate, because human children learn to connect their preexisting intraparietal representation of numerosity with the new arbitrary words and symbols that they are taught. During symbolic calculation, humans quickly access this quantity representation, and they rely on its approximate numerosity code for operations of comparison and approximate calculation. Thus, an evolutionarily ancient representation is put to use for culturally novel symbolic manipulations, including elaborate mathematical ones.[1]

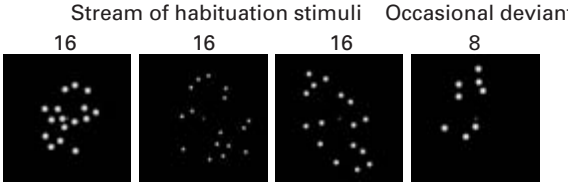
The hypothesis of a reliance on the animal numerosity representation during human symbolic operations leads to several predictions, many of which have been verified. First, human adults, even during symbolic task with Arabic numerals, should show evidence of analog magnitude processing. Indeed, continuous distance effects and Weber's law are characteristically observed in both human and animals in a broad variety of symbolic and nonsymbolic tasks (Barth, Kanwisher, & Spelke, 2003; Cordes, Gelman, Gallistel, & Whalen, 2001; Pinel et al., 2001; Whalen, Gallistel, & Gelman, 1999). This holds even when such effects are deleterious to performance (e.g., when being slower to compare 59 with 65 than 51 with 65, although focusing on the left-most digit would seem sufficient to decide that both are smaller) (Dehaene & Akhavein, 1995; Dehaene, Dupoux, & Mehler, 1990). In fMRI, distance-dependent

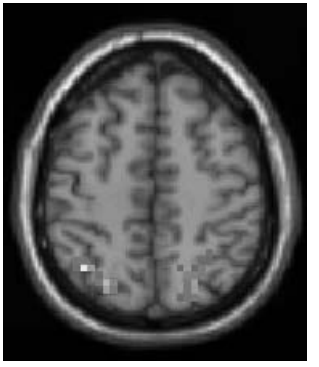
Numerosity detectors (Dehaene & Changeux, 1993)

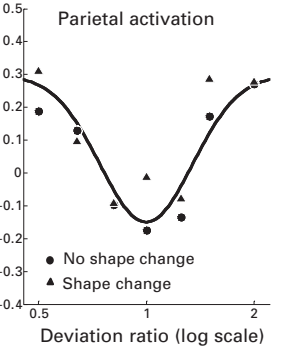


Habituation experiment









numerical priming, conceptually similar to the above numerosity-dependent habituation, has been obtained with masked Arabic and verbal numerals, suggesting that this region is coding jointly for numbers presented in symbolic and non-symbolic form (Naccache & Dehaene, 2001).

A second consequence is that human infants, prior to schooling or even to language acquisition, should exhibit a primitive "number sense" comparable to animals. Indeed, behavioral evidence indicates that infants in the first year of life can process numbers. Although there is some debate about the origins of this competence (Feigenson, Carey, & Hauser, 2002), it arises at least in part from an analog magnitude system similar to the monkey's, capable of dealing with relatively large approximate numerosities (e.g., 8 versus 16) and obeying Weber's law (Brannon, 2002; Xu, 2003; Xu & Spelke, 2000).

Third, early lesions of the HIPS should severely interfere with the development of arithmetic. Indeed, recent neuroimaging studies of children suffering from developmental dyscalculia have revealed demonstrable intraparietal insults that can sometimes be dated to prenatal or perinatal injuries (Isaacs, Edmonds, Lucas, & Gadian, 2001; L. M. Levy, Reis, & Grafman, 1999). My colleagues and I recently showed that a genetic disease, Turner's syndrome, is associated with behavioral, neuroanatomic, and functional activation impairments associated with the intraparietal sulcus (Molko et al., 2003). The existence of such selective impairments in other normally intelligent children supports the view that arithmetic does not emerge solely from a cultural construction process, but requires the integrity of specific brain structures that provide a conceptual foundation for learning.

Cerebral Bases of Reading

The Visual Word Form Area

I now turn to the cerebral bases of another important human cultural invention: reading. Reading even a single word activates a distributed set of brain regions (Fiez & Petersen, 1998), many of which are shared with spoken language processing. Here,

◀ **Figure 8.2**
Evidence for numerosity tuning in the human intraparietal cortex. Dehaene and Changeux's (1993) model of number processing postulated "numerosity-detector" neurons each tuned to an approximate numerosity (*top*), an hypothesis that was recently confirmed by single-neuron recordings in the monkey (see chapter 7, this volume). A habituation design was used to probe the existence of a similar code in humans (*middle*). By repeatedly presenting a fixed numerosity (here 16) and then probing the fMRI response to various deviant numerosities, a tuning curve for numerosity change, here expressed as a Gaussian function of the log ratio of deviant to habituation number, could be observed in the HIPS region (*bottom*) (Piazza et al., 2004). This suggests a similar principle of number tuning in monkeys and humans.

however, I concentrate on one activation site, located in the left occipitotemporal sulcus. There is evidence that this region, which has been termed the "visual word form area" (VWFA), is highly attuned to words in the subject's acquired script. This presents an apparent paradox parallel to the one raised by studies of calculation, inasmuch as there has not been any evolutionary time to evolve a brain area dedicated to reading.

The VWFA is easily identified by collecting fMRI data during short presentations of written words, under passive viewing instructions as well as during active tasks such as semantic classification. Activation is systematically observed in the left occipitotemporal sulcus on the lateral border of the fusiform gyrus (figure 8.3A, plate 16), whether words are contrasted with a fixation control (Dehaene, Le Clec'H, Poline, Le Bihan, & Cohen, 2002) or with presentation of more controlled visual stimuli such as checkerboards (Cohen et al., 2002), pictures of faces, textures, or buildings (Gauthier et al., 2000; Hasson, Levy, Behrmann, Hendler, & Malach, 2002; Puce, Allison, Asgari, Gore, & McCarthy, 1996), pseudo-letters or even random consonant strings (Cohen et al., 2002; Price, Wise, & Frackowiak, 1996). The VWFA can be identified in any single subject (Cohen et al., 2000; Cohen et al., 2002; Dehaene et al., 2002; Gauthier et al., 2000; Puce et al., 1996), allowing quantification of its spatial variability, which appears remarkably low. The standard deviation of its peak coordinates in the Talairach system is about 5 millimeters (Cohen et al., 2000; Dehaene et al., 2002). Furthermore, it occupies a systematically more lateral location relative to the fusiform activation induced by faces (Puce et al., 1996), and falls at a systematic location relative to larger-scale maps of retinotopic and object preference (Hasson, Levy, Behrmann, Hendler, & Malach, 2002).

This reproducible localization is incompatible with many connectionist models of learning to read. Polk and Farah (1998), for instance, presented a model in which distinct regions for numbers and letters emerged through Hebbian learning within an initially unbiased visual layer. In their model, neurons dedicated to letters emerged at a random location and then grouped together to form local letter-sensitive patches. The neuroimaging data suggest that this view cannot be correct. The sites of visual activation during reading are both restricted and highly reproducible, suggesting that there is considerable bias in the underlying cortical tissue prior to learning to read.

Evidence for Functional Specialization and Cultural Impregnation

Three pieces of evidence indicate that the VWFA is functionally specialized for extracting an abstract, invariant representation of letters strings. First, it only activates for visual, not for spoken words (unless the task induces top-down processing; for discussion, see Cohen & Dehaene, 2004). Furthemore, its lesioning leads to pure alexia, a deficit of visual but not spoken word recognition. Indeed, there is good anatomical

convergence between the activation site during reading in normal subjects, and the common lesion site in pure alexia (see figure 8.3A; Cohen et al., 2003). Second, the VWFA is activated by both real words and pronounceable pseudo-words, more than by consonant strings (Cohen et al., 2002; Dehaene et al., 2002; Price et al., 1996). This suggests an intervention at a prelexical stage of processing, and also implies that this area has been changed by learning to read. The selection of letters shapes is an arbitrary cultural convention, and the consonant strings stimuli that we used could have been words in another script. Thus, the stronger response to words than to consonant strings indicates that the VWFA has become attuned to reading in a specific script (Cohen et al., 2002).

Third, the VWFA computes an invariant representation of visual words, one that abstracts away from irrelevant surface variations in the visual stimulus. The VWFA is the first visual area that responds in a nonretinotopic manner, with convergence of activation toward the left hemisphere whether the words are presented left or right of fixation (Cohen et al., 2000). In addition to this spatial invariance, using subliminal priming experiments we demonstrated that the VWFA also computes invariance for upper or lower case (figure 8.3B). When a visible target word is preceded by a short, subliminal presentation of a masked prime word, both response time and fMRI activation are reduced if the target and prime correspond to the same word (Dehaene et al., 2001). Crucially, the VWFA is the only visual area in which this repetition suppression phenomenon is independent of case: it is identical whether the visual stimuli are presented in the same case (e.g., radio/radio) or in a different case (e.g., radio/RADIO). In a recent replication, I showed that such cross-case priming obtains even for words made of letters that are highly dissimilar in upper and lower case, and for which the pairing of upper and lower case is merely a matter of cultural convention (e.g., A and a; Dehaene, et al., 2003). Again, this implies that this brain area is very finely attuned to the specific demands of our reading system.

Possible Precursor of the Visual Word Form Area in Monkeys

Altogether, fMRI studies in humans indicate a rather fine functional tuning of a small, reproducible subpart of the visual system to the demands of visual word recognition, including sensitivity to arbitrary cultural conventions such as variations in case. How can such a specialization arise, although the human brain cannot possibly be predisposed for reading? We can shed some light on this issue by consider the function of this area in other primates, or in human prior to learning to read. In humans, the VWFA belongs to the ventral stream for visual recognition. Indeed, even in word-responsive voxels, responses to pictures or drawings of objects can often be elicited (Hasson et al., 2002). fMRI studies comparing the cortical responses to scrambled versus unscrambled objects in humans and macaques suggest that the higher-level regions of visual ventral cortex in humans may be homologous to the inferotemporal

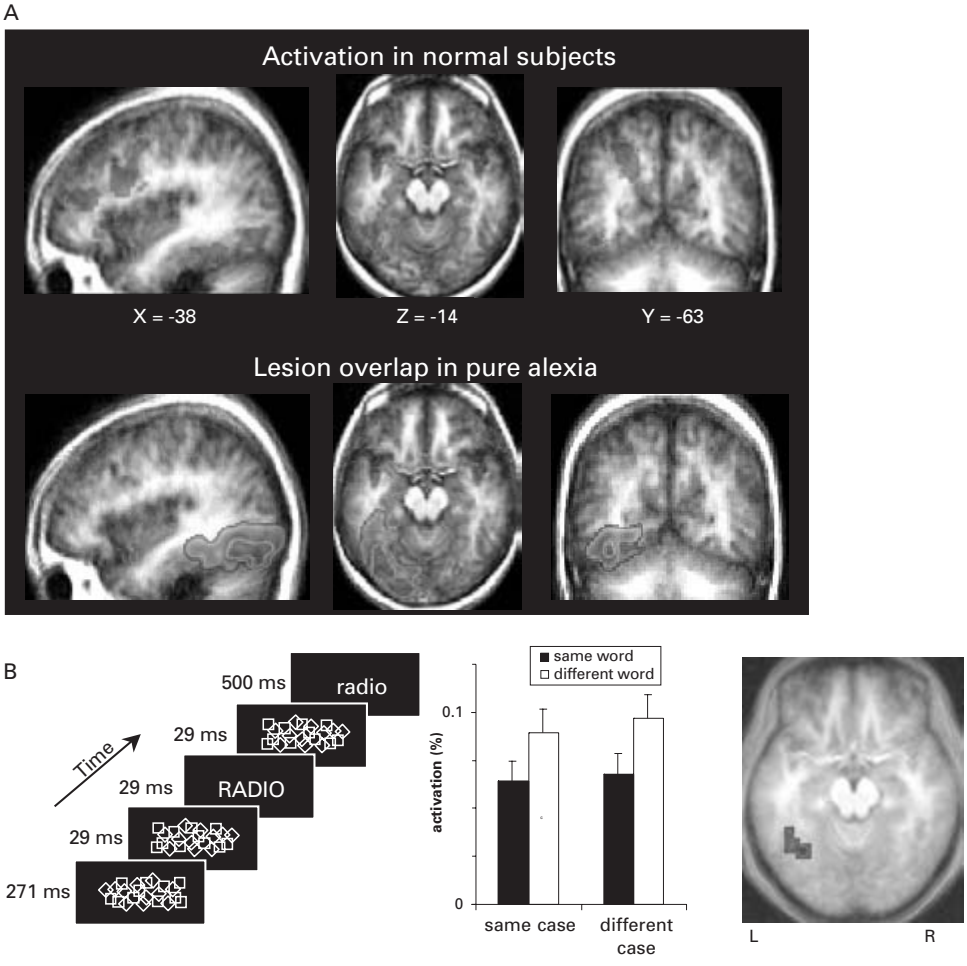


Figure 8.3

Localization and properties of the human visual word form area. The VWFA is defined as a left occipitotemporal region activated in normal subjects in response to visual words more than to control stimuli (here, bright checkerboards); it is also the common lesion site for patients with pure alexia, a severe impairment in visual word identification. (*A* adapted from Cohen et al., 2003) Invariance for case in this region can be demonstrated by using a subliminal priming paradigm, which shows case-independent repetition suppression of the BOLD response only in the left occipitotemporal sulcus. (*B* adapted from Dehaene et al., 2001.) A plot of the latter activation on fiducial (*left*) and flattened (*right*) maps of a human left hemisphere (Caret software, Van Essen et al., 2001) indicates that this area lies just in front of area V8, within inferotemporal cortex involved in higher-level visual recognition. See plate 16 for color version.

C

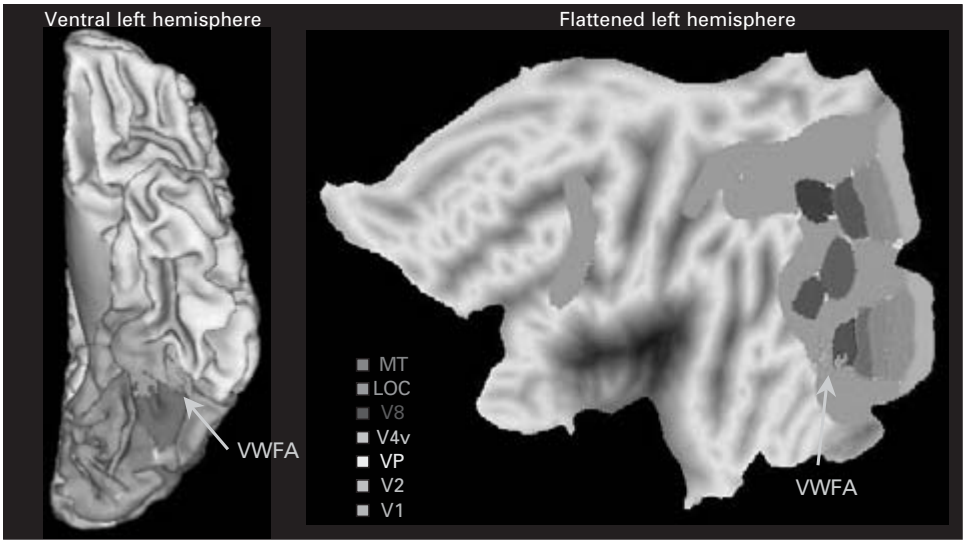


Figure 8.3
(continued)

cortex (area IT) of the macaque (Tootell, Tsao, & Vanduffel, 2003). Furthermore, optical and single-unit recordings indicate that IT neurons possess a high degree of visual invariance (chapter 18, this volume; Ito, Tamura, Fujita, & Tanaka, 1995; Sary, Vogels, & Orban, 1993; Tanaka, 1996). Their receptive fields are vast, often including most or all of the bilateral visual field. They respond preferentially to certain objects, and maintain this preference across a large range of changes in size and retinal location. Some neurons even respond to different views of the same object, for instance the profile and front view of a face, and can learn to respond to arbitrarily related views (Logothetis, 2000; Miyashita, 1988).

These properties suggest that IT neurons are ideally suited to learn to respond to letter, grapheme, and word shapes in a location and case-invariant way. Invariance in visual word recognition may actually result from the intrinsic properties of location and viewpoint invariance found in IT neurons prior to learning to read.

Keiji Tanaka, Manabu Tanifuji and their colleagues have further dissected the selective preferences of IT neurons (see chapter 18). They have observed that, when a neuron responds to a complex object such as the head of a cat, it is often possible to identify a simpler shape to which the neuron is optimally responsive (for instance, a black disk on a white background, similar to an "eye"). IT cortex seems to be composed of a mosaic of such elementary visual detectors (Tanaka, 1996). It is remarkable

that many of those minimal shapes resemble our letters. Some neurons, for instance, fire to two superimposed disks that form a figure of 8, others respond to two bars forming a T, and yet others may respond to an asterisk or a circle. Neurons responsive to these forms may have been selected, during either phylogenesis or ontogenesis, because as an ensemble they provide a repertoire that can represent an immense variety of object shapes. The T shape, for instance, is useful because it frequently signal occlusion of a contour behind some part of the object.

One may therefore speculate that the capacity of this region to learn letter shapes is not an accident. Rather, it derives from the evolutionary and developmental history of IT cortex as a visual recognition system. The minimal shapes that this area can easily represent have been discovered and exploited in our writing systems. In that hypothesis, it is not the human cortex that has evolved for reading—there was not enough evolutionary time and pressure for such an evolution. Rather, writing systems themselves evolved under the constraint of having to remain learnable and easily recognizable by our primate visual system.

Finally, how can one explain the precise location of the VWFA, which is reproducible within a few millimeters across different subjects? There are, in fact, several such examples of precise localization in the visual system. Local preferences for objects, faces, places, and body parts are also fairly reproducible across subjects (chapter 17, this volume; Ishai, Ungerleider, Martin, Schouten, & Haxby, 1999). Rafi Malach and his colleagues suggest that those preferences correspond to fixed locations relative to a large-scale gradient of preference for image excentricity. In human inferotemporal cortex, lateral regions respond preferentially to foveal images, while medial regions prefer parafoveal stimuli (Hasson et al., 2002; Levy, Hasson, Avidan, Hendler, & Malach, 2001; Malach, Levy, & Hasson, 2002). This gradient of excentricity preference cuts across all visual areas of the ventral stream, and may be laid down early on during cortical development, perhaps under the genetic control of an early diffusive "morphogen" substance (Turing's model of morphogenesis). The presence of such an early bias may explain why visual word recognition, which requires high visual accuracy and hence foveation, is systematically located in the lateral inferotemporal cortex. Its lateralization to the left hemisphere might be further explained, similarly, by the presence of privileged connections with multiple language areas of the left hemisphere, particularly the temporal and frontal regions involved in speech comprehension and production.

Learning to Read Changes Human Inferotemporal Cortex

In summary, I speculate that the human brain can learn to read because part of the primate visual ventral object recognition system spontaneously accomplishes operations closely similar to those required in word recognition, and possesses sufficient plasticity to adapt itself to new shapes, including those of letters and words. During

the acquisition of reading, part of this system becomes highly specialized for the visual operations underlying location- and case-invariant word recognition. It occupies a reproducible location within the left occipitotemporal sulcus because the neurons at this location possess intrinsic properties of foveal sensitivity, projection to distant areas in the left hemisphere, and perhaps other undiscovered features that render them most suited for this acquisition. Thus, reading acquisition proceeds by selection and local adaptation of a preexisting neural region, rather than by *de novo* imposition of novel properties onto that region.

In this view, the VWFA should not be considered as a "module" for visual word recognition, but rather as a population of neurons, distributed and overlapping with other populations involved in object recognition, which becomes progressively attuned to the reading process. This view predicts that preference for written words in this region should be relative rather than absolute, and should emerge progressively during learning to read, as the child acquires increasing expertise in word recognition. Indeed, developmental fMRI studies have identified a correlation between VWFA activation and reading skill (Shaywitz et al., 2002). Furthermore, in dyslexic individuals, this region does not respond normally to letters and words (Paulesu et al., 2000; Shaywitz et al., 1998). This reduced activation may not be causally related to dyslexia, but rather may reflect a lack of automatization of word recognition resulting from a primary phonological deficit.

General Principles of Cultural Preemption

The two examples of cultural activities that I have considered, arithmetic and reading, exhibit significant commonalities, but also differences. In both cases, humans learn to attribute meaning to conventional shapes (Arabic digits or the alphabet), and they eventually do so in a highly efficient manner, even subliminally. Furthermore, the brain activations associated with these cultural activities are highly reproducible. Finally, the brain areas involved turn out to have a significantly related function in primate evolution. There is however an important difference between arithmetic and reading. On the one hand, there is a genuine precursor of number knowledge in primate evolution. Intraparietal cortex already seems to be involved in number representation in primates, and the cultural mapping of number symbols onto this representation significantly enhances, but does not radically modify its computational capacity. On the other hand, the evolutionary precursor of the visual word form area did not evolve for reading. It evolved for object recognition, a function significantly different from the mapping of written language onto sound and meaning.

As a generalization of those two examples, I tentatively propose that the human ability to acquire new cultural objects relies on a neuronal "reconversion" or "recycling" process whereby those novel objects invade cortical territories initially devoted

to similar or sufficiently close functions.[2] According to this view, our evolutionary history, and therefore our genetic organization, has created a cerebral architecture that is both constrained and partially plastic, and that delimits a space of learnable cultural objects. New cultural acquisitions are therefore possible only inasmuch as they are able to fit within the preexisting constraints of our brain architecture.

The present hypothesis bears considerable similarity with a classical Darwinian concept which has been called "tinkering" by François Jacob (1977) or "exaptation" by Gould and Vrba (1982)—the reutilization, during phylogenesis, of biological mechanisms for a new function different from the one for which they evolved. In the case of cultural objects, however, this process takes place at a shorter time scale of weeks, months or years, through epigenetic mechanisms that do not require any change in the genome. The terms "reconversion" or "recycling" capture the idea that this process occurs in the lifetime of the individual: each cultural acquisition must find its ecological niche in the human brain, a circuit whose initial role is close enough and whose flexibility is sufficient to be reconverted to this new role.

The terms "reconversion" or "recycling" also make clear that the neuronal tissue that supports cultural learning is not a blank slate, but possesses prior properties (though perhaps only in the form of small biases). Not any kind of object can be made of recycled glass or paper: those materials possess intrinsic physical properties that make them more suitable for some uses than for others. Likewise, each cortical region or network possesses intrinsic properties that are adapted to the function it evolved for, and are only partially modifiable during the cultural acquisition process. Cultural learning in humans may never totally overturn such preexisting biases, but rather changes them minimally as needed. Thus, cultural objects may not be infinitely malleable, and should in fact often reflect intrinsic constraints of the underlying neural networks.

I end by emphasizing three consequences of this view, and examining how they might apply to reading and arithmetic.

Prediction 1: Our genetic envelope should limit the set of learnable cultural objects. Contrary to the view that learning is an open-ended source of unbounded cultural variation, the recycling hypothesis predicts that the human capacity for cultural invention, although extensive, is eventually limited by the envelope of possibilities inherent in our brain circuits. This should lead to a reanalysis of the extent of cultural diversity. There may be a common structure beyond the obvious cultural variations. Seen in this light, writing systems, for instance, appear as relatively invariable: they all use a small repertoire of highly contrasted, basic, foveal shapes; they all map those shapes onto a mixture of sounds and morphemes; and they all take for granted that character size and location are irrelevant (although this invariance does not need to be explicitly taught). Some of these properties may reflect the evolutionary constraints of the cere-

bral circuits that are preempted when we acquire reading. This view also predicts that there should be unlearnable writing systems. Although this has not been tested, it seems likely that computer bar codes, for instance, in which information is encoded in binary form by fine metric cues, would not be learnable by a human.

Prediction 2: Learning difficulty should depend on the distance between the initial function and the new one. It should possible to account for the difficulty of acquiring a new cultural tool based on the amount of transformation that separates the initial, evolutionarily inherited function of the underlying brain circuits and the new, culturally acquired one. The recycling hypothesis predicts that preexisting biases should often speed up the cultural acquisition of novel material. In arithmetic, for instance, the availability of a preverbal analog representation of number magnitude is thought to facilitate the acquisition of Arabic symbols and the counting sequence, because it provides even very young children with an intuitive grasp of the number domain and its basic principles (Dehaene, 1997; Gelman & Gallistel, 1978). In reading, similarly, the properties of size and location invariance that are intrinsic to the visual system are likely to considerably speed up reading acquisition because they provide a stable visual representation of letters to correlate with phonological representations of word sound. The ease or "transparency" of this mapping may then become a crucial determinant of speed and efficiency of learning to read in different languages (e.g., Paulesu et al., 2001).

More generally, the efficiency of education should be greatly enhanced by using teaching strategies that capitalize upon the preexisting representations that young children possess prior to entering school. For instance, finger counting, token counting, and the abacus may provide excellent support for early arithmetic learning, since they rely upon small sets of movable objects whose numerosity is perceivable in infancy, to support the acquisition of more abstract arithmetic computations.

Occasionally, however, some of the child's preexisting cerebral representations may run counter to what needs to be learned. The necessity to unlearn features that were useful in our evolution, but are now counterproductive for the current cultural use of a given brain area, may explain the striking difficulties that some school topics pose to all children. In arithmetic, negative numbers and fractions are good examples of difficult concepts that may go significantly beyond the existing representational capacities of the preverbal primate brain, because they violate basic principles of integer arithmetic (for instance, that adding and multiplying always result in a larger number). Similarly in reading, letters that are mirror images of each other may pose a special challenge for our visual system. Inferotemporal neurons appear to generalize spontaneously across left-right symmetry, preferring the same object whether it is facing left or right (Rollenhagen & Olson, 2000). Contrary to location and size invariance, this invariance across mirror symmetry, although useful in object recognition, may be deleterious for reading as it may lead to confusion of the letters *p* and *q*, or *b* and *d*. This

may explain the peculiar errors that young children make, sometimes writing single letters or even entire words in mirror image without noticing it (Orton, 1925). If my hypothesis of a recyling of the ventral object recognition system for reading is correct, this form of mirror-image generalization needs to be unlearned during the acquisition of reading.

Prediction 3: Cultural learning may reduce the cortical space available for previous abilities. In many cases, cultural learning improves on an existing biological function. For instance, in the arithmetic domain, new symbolic and linguistic representations of numerals become connected to the analog quantity representation. These new connections make quantity information quickly available in a broad variety of multimodal contexts, and they may even improve the precision with which two numbers can be discriminated. In other cases, however, the invasion of an evolutionary older circuit by a new cultural tool may have a measurable cost. This may happen when the old and new functions are incompatible. In such cases, through competition for cortical space, the evolutionary older competence may be reduced or even lost. Learning to read, for instance, may partially displace and reduce object-related activations in the left inferotemporal sulcus. This should have a small cost on the speed or accuracy of visual recognition.

Such a competition effect may not be of much practical import, since it is likely to be detectable only under laboratory conditions. However, it would provide a clear test of the recycling hypothesis. While this prediction does not seem to have been evaluated in the reading domain, it may not be as implausible as it may seem. Indeed, acquisition of visual expertise for cars, which is known to engage inferotemporal cortex within or close to the fusiform face area, was recently shown to interfere with face perception. In comparison to control subjects, experts in car recognition who were asked to memorize cars and faces on alternate trials showed evidence of reduced holistic processing of faces, both in behavioral performance and in the amplitude of the right-hemispheric face-evoked event-related potential (Gauthier, Curran, Curby, & Collins, 2003). If replicated, this result may indicate that the acquisition of car expertise interferes with some components of face recognition processes.

Conclusion

The “neuronal recycling” hypothesis emphasizes that cultural acquisitions must take place within the limited surface and bounded plasticity of the human cortex. The examples of reading and arithmetic indicate that there is more reproducibility in the cortical implementation of those functions than might have been expected based on standard assumptions of large-scale brain plasticity and interindividual variability. A similar degree of anatomical regularity, indicating the existence of significant evolu-

tionary precursors, may exist for other currently understudied cultural domains of human competence such as geometry, algebra, music, and art.

A basic issue remains: Why is it that among primates, only humans invent complex cultural systems such as reading and arithmetic? Various species of primates can be taught to recognize Arabic digits and map them onto quantities (Boysen & Berntson, 1996; Matsuzawa, 1985; Washburn & Rumbaugh, 1991). Thus, the crucial difference may not lie in the capacity to reconvert brain circuits through learning, but in the very ability to *create* new uses for evolutionary older circuits. According to a hypothesis exposed in detail elsewhere (Dehaene, Kerszberg, & Changeux, 1998; Dehaene & Naccache, 2001; Dehaene, Sergent, & Changeux, 2003), the radical expansion of prefrontal cortex and of cortico-cortical connections in our species (see, e.g., chapter 3) may have generated a new ability to mobilize existing processors in a top-down manner within a conscious neuronal workspace. This new circuitry would enable us to tentatively try out new mental syntheses and select them according to their usefulness. Such mental flexibility might have been one of the key factors that lead our ancestors to first try connecting visual recognition processes with phonological and quantity representations, thus making the first crucial steps on the road to reading and arithmetic.

Acknowledgments

I am grateful to Jean-Pierre Changeux for many discussions, and to Marc Hauser and Andreas Nieder for their critical reading of this chapter.

Notes

1. Before we reach a complete theory of number-word acquisition, two issues will have to be clarified through further experimentation. First, there is evidence that another system of "object tracking," able to encode up to three objects, contributes to some but not all numerical tasks in addition to the analog magnitude system. This tracking system is present early on in infancy (Feigenson, Carey, & Hauser, 2002) and exists in other primates (Hauser, Carey, & Hauser, 2000). Its neural basis is currently unknown (though see Piazza, Giacomini, Le Bihan, & Dehaene, 2003; Sathian et al., 1999), as is the exact nature of its contribution to linking symbols and quantities. Second, the human analog quantity representation is probably not passively linked to number symbols, but may be significantly modified in the process—at least in its precision (Weber fraction), but possibly more deeply, for instance in its representation of large numbers and of base 10.

2. The Merriam-Webster dictionary defines the verb "to recycle" as "to pass again through a series of changes or treatments" or "to adapt to a new use." The French term "recyclage" has a slightly different meaning, closer to what I intend to convey. The primary meaning of "recyclage" applies to students or employees and refers to a change in their orientation or to a complementary

formation period designed to adapt them to a new job (the English equivalent might be "retraining" or "reorientation"). It should be clear that my use of the word "recycling" does not imply that the initial function of a given brain area, prior to cultural acquisition, should be considered as garbage, as one referee suggested! I emphasize that cultural reconversion or "neuronal recycling" transforms what was initially a useful function in our evolutionary past into another function that is currently more useful within the present cultural context.

References

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition*, *86*(3), 201–221.

Boysen, S. T., & Berntson, G. G. (1996). Quantity-based interference and symbolic representations in chimpanzees (*Pan trogolodytes*). *Journal of Experimental Psychology: Animal Behavior Processes*, *22*, 76–86.

Brannon, E. M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition*, *83*(3), 223–240.

Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science*, *282*(5389), 746–749.

Chochon, F., Cohen, L., van de Moortele, P. F., & Dehaene, S. (1999). Differential contributions of the left and right inferior parietal lobules to number processing. *Journal of Cognitive Neuroscience*, *11*, 617–630.

Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: The case for the Visual Word Form Area. *NeuroImage*, *22*, 466–476.

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M. A., et al. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, *123*, 291–307.

Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*, *125*(5), 1054–1069.

Cohen, L., Martinaud, O., Lemer, C., Lehéricy, S., Samson, Y., Obadia, M., et al. (2003). Visual word recognition in the left and right hemispheres: Anatomical and functional correlates of peripheral alexias. *Cerebral Cortex*, *13*, 1313–1333.

Cordes, S., Gelman, R., Gallistel, C. R., & Whalen, J. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin and Review*, *8*(4), 698–707.

Courtney, S. M., Petit, L., Maisog, J. M., Ungerleider, L. G., & Haxby, J. V. (1998). An area specialized for spatial working memory in human frontal cortex. *Science*, *279*(5355), 1347–1351.

Dehaene, S. (1997). *The Number Sense*. New York: Oxford University Press.

Dehaene, S., & Akhavein, R. (1995). Attention automaticity and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 314–326.

Dehaene, S., & Changeux, J. P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience*, *5*, 390–407.

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition*, *1*, 83–120.

Dehaene, S., & Cohen, L. (1997). Cerebral pathways for calculation: Double dissociation between rote verbal and quantitative knowledge of arithmetic. *Cortex*, *33*, 219–250.

Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital: Analogical and Symbolic effects in two-digit number comparison. *Journal of Experimental Psychology*: *Human Perception and Performance*, *16*, 626–641.

Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J. B., Le Bihan, D., et al. (2004). Letter binding and invariant recognition of masked words: Behavioral and neuroimaging evidence. *Psychological Science*, *15*, 307–313.

Dehaene, S., Kerszberg, M., & Changeux, J. P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences*, *95*(24), 14529–14534.

Dehaene, S., Le Clec'H, G., Poline, J. B., Le Bihan, D., & Cohen, L. (2002). The visual word form area: A prelexical representation of visual words in the fusiform gyrus. *Neuroreport*, *13*(3), 321–325.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, *79*, 1–37.

Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J. F., Poline, J. B., et al. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, *4*, 752–758.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, *20*, 487–506.

Dehaene, S., Sergent, C., & Changeux, J. P. (2003). A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proceedings of the National Academy of Sciences*, *100*, 8520–8525.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science*, *284*(5416), 970–974.

Duffau, H., Denvil, D., Lopes, M., Gasparini, F., Cohen, L., Capelle, L., et al. (2002). Intraoperative mapping of the cortical areas involved in multiplication and subtraction: An electrostimulation study in a patient with a left parietal glioma. *Journal of Neurology Neurosurgery and Psychiatry*, *73*(6), 733–738.

Eger, E., Sterzer, P., Russ, M. O., Giraud, A. L., & Kleinschmidt, A. (2003). A supramodal number representation in human intraparietal cortex. *Neuron*, *37*(4), 719–725.

Feigenson, L., Carey, S., & Hauser, M. (2002). The representations underlying infants' choice of more: object files versus analog magnitudes. *Psychological Science*, *13*(2), 150–156.

Fias, W., Lammertyn, J., Reynvoet, B., Dupont, P., & Orban, G. A. (2003). Parietal representation of symbolic and nonsymbolic magnitude. *Journal of Cognitive Neuroscience*, *15*(1), 47–56.

Fiez, J. A., & Petersen, S. E. (1998). Neuroimaging studies of word reading. *Proceedings of the National Academy of Sciences*, *95*(3), 914–921.

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience*, *6*(4), 428–432.

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience*, *12*(3), 495–504.

Gelman, R., & Gallistel, C. R. (1978). *The Child's Understanding of Number*. Cambridge, MA.: Harvard University Press.

Gould, S. J., & Vrba, E. S. (1982). Exaptation: A missing term in the science of form. *Paleobiology*, *8*, 4–15.

Gruber, O., Indefrey, P., Steinmetz, H., & Kleinschmidt, A. (2001). Dissociating neural correlates of cognitive components in mental calculation. *Cereb Cortex*, *11*(4), 350–359.

Hasson, U., Levy, I., Behrmann, M., Hendler, T., & Malach, R. (2002). Eccentricity bias as an organizing principle for human high-order object areas. *Neuron*, *34*(3), 479–490.

Hauser, M. D., Carey, S., & Hauser, L. B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proceedings of the Royal Society of London B Biological Science*, *267*(1445), 829–833.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, *298*(5598), 1569–1579.

Hauser, M. D., Dehaene, S., Dehaene-Lambertz, G., & Patalano, A. L. (2002). Spontaneous number discrimination of multi-format auditory stimuli in cotton-top tamarins (*Saguinus oedipus*). *Cognition*, *86*(2), B23–B32.

Isaacs, E. B., Edmonds, C. J., Lucas, A., & Gadian, D. G. (2001). Calculation difficulties in children of very low birthweight: A neural correlate. *Brain*, *124*(9), 1701–1707.

Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L., & Haxby, J. V. (1999). Distributed representation of objects in the human ventral visual pathway. *Proceedings of the National Academy of Sciences*, *96*(16), 9379–9384.

Ito, M., Tamura, H., Fujita, I., & Tanaka, K. (1995). Size and position invariance of neuronal responses in monkey inferotemporal cortex. *Journal of Neurophysiology*, *73*, 218–226.

Jacob, F. (1977). Evolution and tinkering. *Science*, *196*(4295), 1161–1166.

Lee, K. M. (2000). Cortical areas differentially involved in multiplication and subtraction: A functional magnetic resonance imaging study and correlation with a case of selective acalculia. *Annals of Neurology*, *48*, 657–661.

Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object areas. *Nature Neuroscience*, *4*(5), 533–539.

Levy, L. M., Reis, I. L., & Grafman, J. (1999). Metabolic abnormalities detected by H-MRS in dyscalculia and dysgraphia. *Neurology*, *53*, 639–641.

Logothetis, N. (2000). Object recognition: Holistic representations in the monkey brain. *Spatial Vision*, *13*(2–3), 165–178.

Malach, R., Levy, I., & Hasson, U. (2002). The topography of high-order human object areas. *Trends in Cognitive Science*, *6*(4), 176–184.

Matsuzawa, T. (1985). Use of numbers by a chimpanzee. *Nature*, *315*, 57–59.

McComb, K., Packer, C., & Pusey, A. (1994). Roaring and numerical assessment in contests between groups of female lions, Panthera leo. *Animal Behaviour*, *47*, 379–387.

Menon, V., Rivera, S. M., White, C. D., Glover, G. H., & Reiss, A. L. (2000). Dissociating prefrontal and parietal cortex activation during arithmetic processing. *Neuroimage*, *12*(4), 357–365.

Miyashita, Y. (1988). Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature*, *335*(6193), 817–820.

Molko, N., Cohen, L., Bruandet, M., Mangin, J. F., Cachia, A., Le Bihan, D., et al. (2003). Functional and structural abnormalities of the intraparietal sulcus in a genetic developmental dyscalculia. *Neuron*, 40, 847–858.

Naccache, L., & Dehaene, S. (2001). The priming method: Imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cerebral Cortex*, *11*(10), 966–974.

Nakahara, K., Hayashi, T., Konishi, S., & Miyashita, Y. (2002). Functional MRI of macaque monkeys performing a cognitive set-shifting task. *Science*, *295*(5559), 1532–1536.

Nieder, A., Freedman, D. J., & Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science*, *297*(5587), 1708–1711.

Nieder, A., & Miller, E. K. (2003). Coding of cognitive magnitude: Compressed scaling of numerical information in the primate prefrontal cortex. *Neuron*, *37*(1), 149–157.

Orton, S. T. (1925). "Word-Blindness" in School Children. *Archives of Neurology and Psychiatry*, *14*, 581–615.

Paulesu, E., Demonet, J. F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., et al. (2001). Dyslexia: cultural diversity and biological unity. *Science*, *291*(5511), 2165–2167.

Paulesu, E., McCrory, E., Fazio, F., Menoncello, L., Brunswick, N., Cappa, S. F., et al. (2000). A cultural effect on brain function. *Nature Neuroscience*, *3*(1), 91–96.

Pesenti, M., Thioux, M., Seron, X., & De Volder, A. (2000). Neuroanatomical substrates of arabic number processing, numerical comparison, and simple addition: A PET study. *Journal of Cognitive Neuroscience*, *12*(3), 461–479.

Piazza, M., Giacomini, E., Le Bihan, D., & Dehaene, S. (2003). Single-trial classification of parallel pre-attentive and serial attentive processes using functional magnetic resonance imaging. *Proceedings of the Royal Society London B Biological Sciences*, *270*(1521), 1237–1245.

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron*, *44*, 547–555.

Pinel, P., Dehaene, S., Riviere, D., & LeBihan, D. (2001). Modulation of parietal activation by semantic distance in a number comparison task. *Neuroimage*, *14*(5), 1013–1026.

Polk, T. A., & Farah, M. J. (1998). The neural development and organization of letter recognition: Evidence from functional neuroimaging, computational modeling, and behavioral studies. *Proceedings of the National Academy of Sciences*, *95*(3), 847–852.

Price, C. J., Wise, R. J. S., & Frackowiak, R. S. J. (1996). Demonstrating the implicit processing of visually presented words and pseudowords. *Cerebral Cortex*, *6*, 62–70.

Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: A functional magnetic resonance imaging study. *Journal of Neuroscience*, *16*, 5205–5215.

Quartz, S. R., & Sejnowski, T. J. (1997). The neural basis of cognitive development: A constructivist manifesto. *Behavior and Brain Science*, *20*(4), 537–556; discussion, 556–596.

Rickard, T. C., Romero, S. G., Basso, G., Wharton, C., Flitman, S., & Grafman, J. (2000). The calculating brain: An fMRI study. *Neuropsychologia*, *38*(3), 325–335.

Rollenhagen, J. E., & Olson, C. R. (2000). Mirror-image confusion in single neurons of the macaque inferotemporal cortex. *Science*, *287*(5457), 1506–1508.

Sary, G., Vogels, R., & Orban, G. A. (1993). Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science*, *260*(5110), 995–997.

Sathian, K., Simon, T. J., Peterson, S., Patel, G. A., Hoffman, J. M., & Grafton, S. T. (1999). Neural evidence linking visual object enumeration and attention. *Journal of Cognitive Neuroscience*, *11*(1), 36–51.

Sawamura, H., Shima, K., & Tanji, J. (2002). Numerical representation for action in the parietal cortex of the monkey. *Nature*, *415*(6874), 918–922.

Shaywitz, S. E., Shaywitz, B. A., Pugh, K. R., Fulbright, R. K., Constable, R. T., Mencl, W. E., et al. (1998). Functional disruption in the organization of the brain for reading in dyslexia. *Proceedings of the National Academy of Sciences*, *95*(5), 2636–2641.

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarski, P., et al. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biological Psychiatry*, *52*(2), 101–110.

Simon, O., Mangin, J. F., Cohen, L., Le Bihan, D., & Dehaene, S. (2002). Topographical layout of hand, eye, calculation, and language-related areas in the human parietal lobe. *Neuron*, *33*(3), 475–487.

Simon, T. (1999). Finding the foundations of numerical thinking in a brain without numbers. *Trends in Cognitive Science*, *3*, 363–365.

Stanescu-Cosson, R., Pinel, P., van de Moortele, P.-F., Le Bihan, D., Cohen, L., & Dehaene, S. (2000). Cerebral bases of calculation processes: Impact of number size on the cerebral circuits for exact and approximate calculation. *Brain*, *123*, 2240–2255.

Takayama, Y., Sugishita, M., Akiguchi, I., & Kimura, J. (1994). Isolated acalculia due to left parietal lesion. *Archives of Neurology*, *51*, 286–291.

Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, *19*, 109–139.

Tootell, R. B., Tsao, D., & Vanduffel, W. (2003). Neuroimaging weighs in: Humans meet macaques in "Primate" visual cortex. *Journal of Neuroscience*, *23*(10), 3981–3989.

Van Essen, D. C., Drury, H. A., Dickson, J., Harwell, J., Hanlon, D., & Anderson, C. H. (2001). An integrated software suite for surface-based analyses of cerebral cortex. *Journal of the American Medical Informatics Association*, *8*(5), 443–459.

Washburn, D. A., & Rumbaugh, D. M. (1991). Ordinal judgments of numerical symbols by macaques (*Macaca mulatta*). *Psychological Science*, *2*, 190–193.

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Non-verbal counting in humans: The psychophysics of number representation. *Psychological Science*, *10*, 130–137.

Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition*, *89*(1), B15–25.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), B1–B11.

Zago, L., Pesenti, M., Mellet, E., Crivello, F., Mazoyer, B., & Tzourio-Mazoyer, N. (2001). Neural correlates of simple and complex mental calculation. *Neuroimage*, *13*(2), 314–327.

9 Cooperative Brains: Psychological Constraints on the Evolution of Altruism

Jeffrey R. Stevens and Marc D. Hauser

Imagine an individual called "hunter" that expends a good deal of energy to capture a gazelle. As the hunter is consuming his small prey, a second individual called "recipient" approaches and begins feeding peacefully alongside the hunter. A few weeks later the roles reverse, such that the previous recipient has now captured a gazelle, and the previous hunter is taking advantage of the recipient's hard work. Could the hunter and recipient be Maasai warriors? Is it equally likely that they are common chimpanzees, African lions, or Nile crocodiles? All of these species hunt gazelle and live in groups, so why would this scenario apply to some species more appropriately than to others? The answer lies in the costs and benefits associated with sharing food with non-kin. Assuming that one individual can consume the entire gazelle, sharing food with the recipient constitutes an altruistic act—the hunter accepts a fitness cost (reduction in food intake) while increasing the fitness of another (increasing the intake of the recipient).

Reciprocal Altruism: Cooperation via Turn Taking

Here we focus on a form of altruistic cooperation with interesting psychological implications: reciprocal altruism (or reciprocity)—the alternation of donor and recipient roles in repeated altruistic interactions. When Trivers (1971) first introduced the concept of reciprocal altruism, he outlined necessary prerequisites such as a large benefit to the recipient and a small cost to the donor, many opportunities for cooperative interactions, and the ability to detect cheaters. Humans appear to satisfy these requirements quite nicely.[1] Intuitively, reciprocity seems fundamental to human social interactions, and experiments confirm this intuition, demonstrating its prevalence across different economic contexts (Fehr & Fischbacher, 2003; Fehr & Henrich, 2003; Gintis, Bowles, Boyd, & Fehr, 2003; McCabe, 2003), as well as across cultures (Henrich et al., 2001). In fact, some argue that it is so integral to human society that we have evolved specialized cognitive mechanisms to facilitate reciprocal interactions

including especially, the detection of cheaters (Cosmides & Tooby, 1992) and punishment (Fehr & Gächter, 2000).

Given that reciprocity is common in humans, emerges fairly early in development (Harbaugh, Krause, & Liday, 2002; Harbaugh, Krause, Liday, & Vesterlund, 2003), and that the prerequisites appear trivial, should we expect to see it in nonhuman animals? In this chapter, we address this question by developing the following argument. First, we argue that the prerequisites for reciprocal altruism have been underestimated. A careful dissection reveals a host of underlying mechanisms that may be necessary for both initiating a reciprocal relationship and for maintaining it over the long haul. Second, we argue that some of the essential psychological ingredients for reciprocation include numerical quantification, time estimation, delayed gratification, detection and punishment of cheaters, analysis and recall of reputation, and inhibitory control. For example, reciprocal altruism requires inhibitory control in order to suppress the temptation to cheat: once B has received from A, B must inhibit the temptation to defect in order to return the favor and maintain a stable cooperative relationship. Reciprocal altruism also requires, by definition, patience: if A gives to B, A must wait some period of time before B returns the favor. And reciprocal altruism may require quantification: A and B may quantify the resources exchanged, potentially across different currencies, in order to evaluate whether the exchange was fair. If reciprocity is, indeed, as cognitively complex as we suggest, then we must anchor our theoretical predictions about adaptive function in realistic constraints imposed by neural and psychological design features. Third, because of these limitations, we predict that reciprocal altruism will be rare in the animal kingdom, and when it appears, will represent a relatively minor force in the evolution of social organizations. Returning to our opening paragraph, although reciprocal altruism represents a theoretical solution to the problem of altruism between unrelated individuals, does the crocodile, lion, or chimpanzee have the cognitive wherewithal for reciprocity?

We begin by critically discussing the empirical evidence for reciprocity in animals. For some, we will appear highly critical. We feel this level of analysis is necessary in order to show what we know about cooperation in animals and what we have yet to learn; hopefully, the fact that we are equally critical of our own work as that of others, will make the exercise seem fair. This discussion leads to the conclusion that animals can maintain stable cooperation for mutual, simultaneous benefits, but rarely if ever sustain stable, reciprocally altruistic relationships that entail delayed benefits. We next turn to an explanation of this conclusion, focusing on a suite of cognitive constraints. Finally, we turn to a brief discussion of some of the neurophysiological substrates that might support reciprocation in humans, and use this evidence to speculate about the neural correlates of cooperation in animals.

Reciprocal Altruism: Theoretical Concerns

After Trivers's initial investigation, the concept of reciprocity remained all but untouched until Axelrod and Hamilton's (1981) description of "tit-for-tat" as a possible reciprocal strategy that allows for stable cooperation in the Prisoner's Dilemma (Flood, 1958; Rapoport & Chammah, 1965). In the Prisoner's Dilemma, two individuals each have the opportunity to cooperate with or defect against each other, resulting in four possible fitness payoffs for each player (figure 9.1). Mutual cooperation results in a moderate reward (*R*), but mutual defection leads to very low payoffs for both players (*P*). When one cooperates and the other defects, the defector receives the largest possible reward (*T*) and the cooperator receives the smallest possible reward (*S*). Therefore, the optimum strategy (Nash equilibrium) for playing a single-shot game is to defect, because defection results in a higher payoff than cooperation regardless of the opponent's choice. Axelrod and Hamilton, however, suggested that stable cooperation can emerge if the game is played repeatedly, the opening move is nice (cooperative), and from that point on, each player copies the other's moves. This winning strategy is a version of reciprocity called tit-for-tat. Following this analysis, a flood of theoretical investigations emerged, some confirming the efficacy of reciprocity, others providing alternative strategies that maintain cooperation (reviewed in Dugatkin, 1997).

Empirical Evidence for Reciprocal Behavior

Following the deluge of theory, reciprocity was invoked to explain many instances of animal cooperation. Here we describe case studies that examined putative reciprocal situations, including blood sharing in vampire bats, cooperative games in blue jays, and food exchange in capuchin monkeys and tamarins. Because of space constraints, we leave out the many other interesting cases that have been described, including the

		Against player 2:	
		Cooperate	Defect
Payoff to Player 1:	Cooperate	*R*	*S*
	Defect	*T*	*P*

Figure 9.1

Fitness payoffs in Prisoner's Dilemma game—payoffs *R*, *T*, *P*, and *S* represent payoffs to the row player against the column player (e.g., *T* is the payoff to a defector playing against a cooperator). To qualify as a Prisoner's Dilemma, $T > R > P > S$. Therefore, regardless of the opponent's choice, defection results in larger payoffs, but mutual cooperation is more profitable than mutual defection.

exchange of grooming for alliance support in vervet monkeys (Seyfarth & Cheney, 1984) and reciprocal grooming in antelope and many primates (Hart & Hart, 1992; Muroyama, 1991).

Blood Sharing in Vampire Bats

Vampire bats can live for almost 20 years, spending much of their time in large, stable social groups where there are multiple opportunities to interact with the same individuals. A vampire bat's survival depends critically on the consumption of blood. If an individual goes for more than 60 hours without a blood meal, it dies. On any given day, therefore, an individual must either obtain its own meal or convince another bat to regurgitate some of its undigested blood. This suite of attributes makes vampire bats ideal subjects for studies of reciprocal altruism (Wilkinson, 1984).

Wilkinson (1984) observed more than 100 regurgitations in a wild population of vampire bats. Because blood is valuable, giving it up represents a cost—an act of altruism. Of the cases observed, most were between mother and infant. These were not examined in any detail because there's no puzzle: regurgitating to your offspring makes sense since you share half of your genes with them; there is no expectation of reciprocation here. Of the remaining regurgitations, only 20 percent were between more distant relatives or nonrelatives, and of these, most occurred among bats that frequently spent time together. However, since many of these individuals were genetic relatives (half at the level of grandparent-grandchild), it seems that regurgitation is largely motivated by kinship, with an extremely small proportion of cases among genetically unrelated bats. Nonetheless, given that some regurgitations were delivered to non-kin, these cases require some explanation. There are two possibilities: either some bats made mistakes, failing to recognize their kin and thus accidentally giving blood to non-kin, or they purposefully gave blood to non-kin with the expectation that they would receive blood back in the future.

To better understand what motivates regurgitations among non-kin, and to clarify whether giving is contingent upon receiving, Wilkinson (1984) conducted a simple experiment with eight unrelated vampire bats. Over many days, he removed one bat from the colony before feeding while providing the other bats with two hours of access to blood. He then returned the now starving bat to the blood-satiated bats. The pattern of blood sharing was clear: individuals regurgitated blood to those who had regurgitated to them in the past.

There are four reasons why we want to express caution in accepting the vampire bat case as evidence of reciprocal altruism, even though many authors have trumpeted these observations as some of the best evidence to date (Dugatkin, 1997; Hauser, 2000). One: the number of naturally observed cases is small and can be explained as errors of recognition as opposed to reciprocation among non-kin. Though regurgitations are given to unrelated animals, these are infrequent, and there is no evidence that indi-

viduals recognize the recipients as non-kin as opposed to kin. Wilkinson did not conduct any tests to show that bats recognize their kin, and if so, to what degree of relatedness. The consequence of contingent regurgitation may benefit non-kin, but the payoffs and mechanisms may have evolved for kin, occurring among non-kin as a by-product. Two: the number of experimental cases is also small, and might reflect an artificial outcome, an exchange that is more important in captivity when bats have less certainty with respect to their next meal. Three: even if we accept these few cases, it is not at all clear whether reciprocal altruism among non-kin plays a significant or trivial role in individual survival. The fact that individuals need blood to survive is clear. Whether or not they depend upon reciprocation with non-kin to survive is a different issue. It may well be that individuals would survive fine without it, relying on their own skills, and the good nature of their relatives. Four: only one study has ever attempted to replicate these findings, even though 20 years has elapsed since their original publication. Denault and McFarlane (1995) observed regurgitations among vampire bats, but the degree of relatedness was close to the level of grandparent-grandchild, thereby allowing kinship to account for the pattern of altruistic behavior.

Experimental Games in Blue Jays

A second approach to testing for reciprocal altruism in animals comes from work on captive blue jays trained to peck keys in one of two classical economic games. In most instances of cooperative behavior in animals, the fitness costs and benefits of cooperation remain unclear. This ambiguity makes it difficult to determine whether the animals face a true Prisoner's Dilemma or whether some other benefits to cooperation actually reduce the temptation to cheat. To circumvent these uncertainties, Stephens and colleagues (Clements & Stephens, 1995; Stephens, McLinn, & Stevens, 2002; Stevens & Stephens, 2004) have devised ecologically artificial but economically relevant situations in which blue jays play cooperative games. Their paradigm involves placing pairs of jays in adjacent operant chambers in which each individual simultaneously chooses to cooperate or defect by pecking a key (figure 9.2A). Once both subjects pecked a key, each received a specific number of pellets associated with the game's payoff matrix. For example, in a Prisoner's Dilemma matrix, when both cooperated each received three pellets (*R*), when both defect each received one pellet (*P*), and when one cooperated and the other defected, the defector received five pellets (*T*) and the cooperator received no pellets (*S*). In the second game, called "Mutualism," the payoffs for mutual cooperation were higher than for all other possibilities, and the payoffs for mutual defection were lower than for one player defecting and the other cooperating. When the jays played a Prisoner's Dilemma game, they rapidly defected. In contrast, when the jays switched to a game of Mutualism they not only cooperated but maintained this pattern over many days. That jays switch strategies as a function

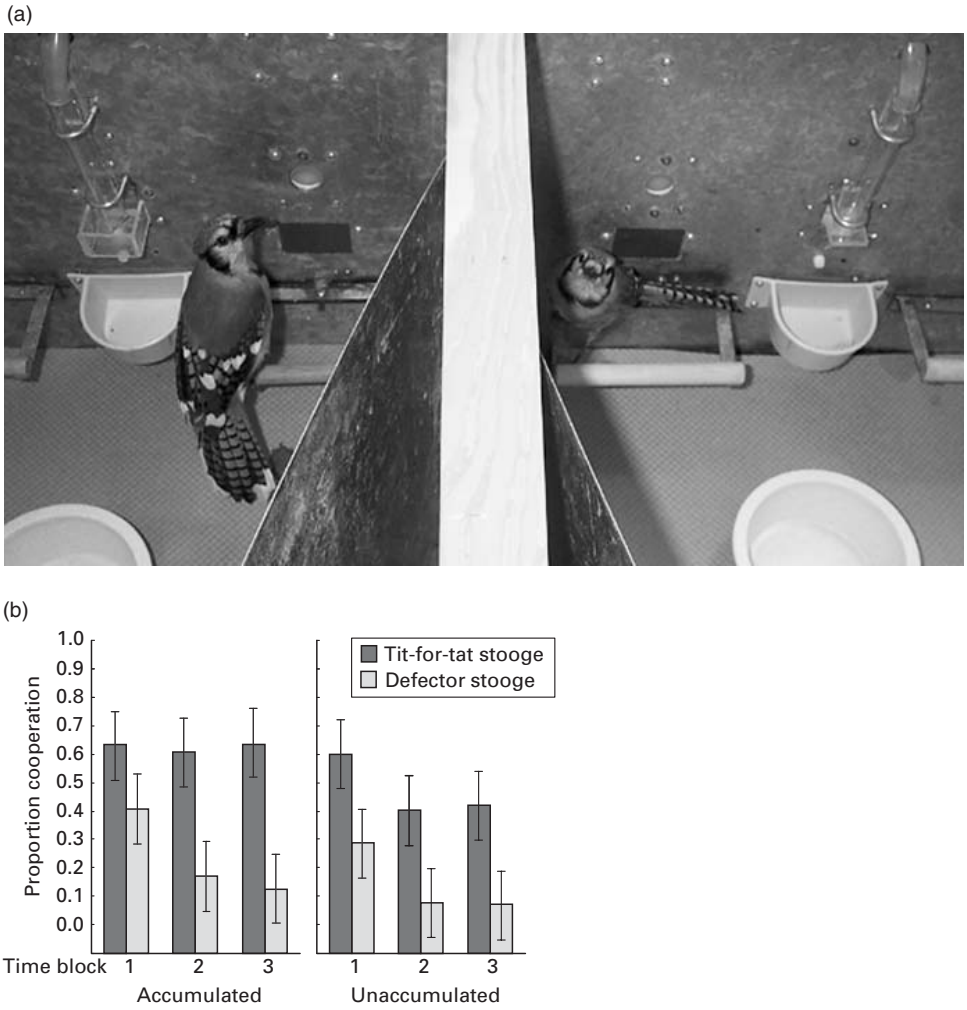


Figure 9.2

Stephens et al. (2002) tested cooperation in an operant experiment with blue jays (*Cyanocitta cristata*) playing Prisoner's Dilemma games. (*A*) Freely behaving subjects played against either tit-for-tat or defector stooges and food rewards were either dispensed after every trial or accumulated over four trials. (*B*) Whereas cooperation decreased over time in all other conditions, cooperation was maintained when subjects played against a tit-for-tat stooge and when food accumulated. This suggests that cooperation depends on opponent strategy as well as temporal discounting. (Adapted from Stephens et al., 2002 with permission from the American Association for the Advancement of Science.)

of the game played shows that their responses are contingent upon the payoffs associated with each game.

Clements and Stephens's results show how cooperation depends upon the relative costs and benefits of different strategies. When cooperation yields the highest possible payoff, jays do not defect. Mutual cooperation is the only reasonable option. In contrast, when there is a temptation to defect, as defined by the Prisoner's Dilemma, then jays are incapable of maintaining a cooperative relationship.

To determine if other conditions might enable cooperation among jays in the Prisoner's Dilemma, Stephens and colleagues ran a second experiment (Stephens, McLinn & Stevens, 2002), this time targeting a potential constraint on the evolution and stability of reciprocal altruism: the temptation to take an immediate benefit outweighs the benefits of waiting for a larger payoff. Numerous studies of animals and humans, discussed more fully below, reveal that waiting for a payoff devalues this item's worth. A small payoff now is better than a large payoff later. This trade-off between time and value is called *discounting*, and is a central idea in economic models of choice. In the original jay work, pecking brought an immediate payoff of some amount. In the new study, payoffs accumulated. To obtain food, each pair of jays had to play several rounds with their partner before obtaining the payoffs, thereby removing the immediate temptation. In addition, the jays played against a partner that either always defected or played tit-for-tat. The jays cooperated only when the food rewards accumulated and when playing against a tit-for-tat partner (figure 9.2B). They solved the repeated Prisoner's Dilemma.

Clements and Stephens concluded their original paper on jays as follows: "[T]here is no empirical evidence of non-kin cooperation in a situation, natural or contrived, where the payoffs are known to conform to a Prisoner's Dilemma" (p. 533). The follow-up studies with jays led Stephens and colleagues to a different conclusion, but one that is consistent with the idea that animals are incapable of maintaining reciprocal relationships under natural conditions: "Our work suggests that the timing of benefits can be the difference between stable cooperation and cooperation that erodes to mutual defection" (p. 2218). The authors also point out that "the experimental machinations required to stabilize cooperation . . . are special" (p. 2218). In other words, nature may never provide animals with the right conditions for reciprocally stable relationships.

Food Sharing in Capuchins

A third example comes from a social primate—the capuchin monkey. Capuchins live in multimale, multifemale social groups with a polygynous mating system. They are a highly dexterous species, and in the wild, hunt in groups and often share food. De Waal attempted to capitalize on their apparent social intelligence by conducting a series of experiments on cooperation. In the first experiment (de Waal & Berger, 2000),

female capuchins had to work for food, either on their own or with another unrelated individual. The task was simple: pull a rod to bring a cup of food within reach. When there were two capuchins and therefore two rods, each individual had to pull at the same time in order to bring the cups within reach. When the experimenter placed food in both cups, both capuchins pulled. Although their joint action is cooperative, it can more readily be explained as selfish, with each individual pulling for herself. When the experimenter placed food in only one cup, the individual lined up with the food almost always pulled whereas the other individual pulled less frequently. Importantly, however, when the player facing an empty cup pulled, she was more likely to obtain food from the other capuchin than when she failed to help. Individuals with access to the food cup rarely handed food to helpers. Instead, they approached helpers and allowed them to grab pieces of food through the partition as it fell to the ground.

Allowing another individual to *take* food is psychologically different from *giving* food. For one, tolerated taking is more difficult to interpret with respect to the motivations or intentions of the possessor. On some occasions, perhaps the possessor did not intend to have food taken but was simply not swift enough to stop the action. Giving, in contrast, is clear-cut, and represents a cost of physical exertion as well as a reduction in one's own resources. In almost 10,000 observations of food transfer, less than 1 percent involved giving. Nonetheless, this experiment shows that capuchins are more likely to tolerate food-taking by an individual who helped them pull in the past.

To explore the contingency part of the capuchin's interactions, as well as the role of food quality, de Waal and colleagues ran other experiments (de Waal, 1997). Individuals were more likely to tolerate food taking when lower quality food items were at stake. Among female-female pairs, individual A was more likely to allow individual B to take food if on the previous run, individual B allowed A to take food. This relationship or correlation accounted for less than 10 percent of the variation in behavior, suggesting that many other factors influence whether or not two females tolerate food-taking; for example, since de Waal did not observe the players before or after the game, we do not know if tolerated food-taking was *repaid* in some other currency such as grooming. Moreover, if two males or a male and female played this game, then tolerated food-taking was not at all contingent on prior food-taking.

A second complication associated with the capuchin studies is that the analyses focus on tolerated food-taking, independently of how much food was taken. Although de Waal reports that the food possessor generally "ate the lion's share," it might well be important to know how much food the nonfood possessor obtained in each of the conditions, as opposed to whether or not it received any food. One might imagine, for instance, that the amount of food taken depends on its quality, on the amount obtained in previous runs, on the food possessor's hunger level, expectations about the number of games to be played in the future, and so forth.

De Waal's work shows that capuchins will tolerate food taking from others, and that this behavior has something to do with the help received on the rod-pulling task. Capuchins clearly cooperate. There are, however, three reasons why we believe that the capuchin work falls short of the required evidence for reciprocation. One: although there is some evidence for reciprocated food exchange, it happens infrequently and is restricted to female-female pairs. Two: Female-female cooperation in captivity may be the by-product of kin selection in nature. Although the females in de Waal's groups were genetically unrelated, in nature, females living in a social group are typically kin. Among most primate species, including capuchins, females stay in their natal groups for life, whereas males emigrate; this leads to groups consisting of closely related females and distantly related males. Reciprocation among female-female pairs could, therefore, be an artificial by-product of selection for kin interactions. It is unclear whether it plays any role in natural groups of capuchins, and if it does, whether it is dwarfed by cooperation among kin. Three: because there is little cost to pulling the rod, and food exchange occurs most frequently when food quality is poor (costs of exchange are low), it is not clear that this task involves altruistic actions; neither the pulling by the helper, nor the tolerated taking of low quality food by the owner, are costly. These three points lead, we suggest, to the conclusion that reciprocal altruism is a weak force in capuchin social relationships.

Food-Giving in Tamarins

A final example comes from another social primate—the cotton-top tamarin. Unlike capuchins that live in large social groups, characterized by polygamous mating behavior, tamarins live in small groups characterized by monogamy. Within groups, which consist of the breeding pair and typically, one to two generations of offspring, older offspring help rear the younger ones. Part of the help comes in the form of food sharing (Feistner & Price, 1990; Roush & Snowdon, 2001). To explore the possibility of reciprocal altruism in tamarins, Hauser and colleagues designed a series of experiments focused on the problem of food exchange (Hauser, Chen, Chen, & Chuang, 2003). Each experiment set up a game between unrelated tamarins in which one animal—the actor—could pull a tool to give food to an unrelated recipient without getting any food for self; pulling the tool is thus altruistic (figure 9.3). Why would unrelated tamarins give each other food?

In the first test, an experimenter trained two tamarin stooges to function as players with diametrically opposite roles: one acted as a unilateral altruist, always pulling the tool to give food to its partner, and the other acted as a unilateral defector, never pulling the tool. The reason for training was simple: if tamarins give food to others based on previous acts of kindness, then they should give most to the altruist and least or nothing to the defector. Supporting this prediction, tamarins pulled the tool most often for the altruist and infrequently for the defector. This shows two things:

(a)

(b)

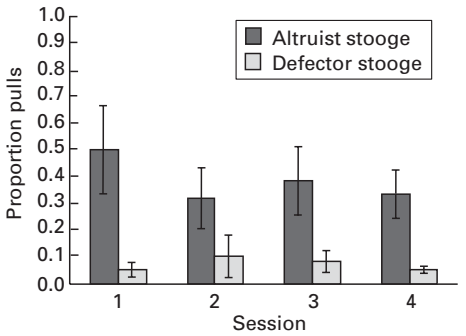


Figure 9.3

Hauser et al. (2003) tested an altruism game in cotton-top tamarins (*Saguinus oedipus*). (*A*) Tamarins had the option to pull a tool, thereby giving food to an unrelated tamarin but receiving nothing in return. Subjects played against an altruist stooge that always pulled or a defector stooges that never pulled. (*B*) The tamarins consistently pulled more often for the altruist stooge than the defector stooge, suggesting that their propensity to cooperate is contingent on their partner's behavior. (Redrawn from Hauser et al., 2003 with permission from the Royal Society of London.)

tamarins give food to unrelated others, and do so based on contingent acts of giving in the past. Is this reciprocal altruism? Not yet. Perhaps tamarins feel more generous when they eat more? When the altruist plays, she gives food on every trial. Her partner must feel good. When a tamarin feels good, it is more likely to pull the tool and give food back. What looks like reciprocation based on an altruistic act of food giving is actually a dumb by-product of feeling good—feeling sated.

To test the "feel good" explanation, Hauser and colleagues ran other experiments, this time leaving the stooges out, and using untrained pairs instead. In one game, player A could pull the tool to obtain a piece of food for self and simultaneously deliver three pieces to player B. On the next trial, player B could pull the tool to give two pieces to player A, but obtain nothing for self. Given these payoffs, reciprocal pulling would pay as each player would obtain three pieces of food after a complete round. Animals in the player A role should always pull out of selfish interest to get food; they did. If feeling good motivates giving food, then player B should cooperate and pull because player A always gives them food—player A looks like a unilateral altruist from one perspective. If player B notices, however, that player A is pulling out of selfish interest—nothing like the unilateral altruist—then player B shouldn't cooperate. Animals in the player B role don't cooperate. Feeling good isn't enough to set reciprocation in motion. And for food giving to count, it can't be an accidental by-product of selfish behavior. It must be given with an altruistic intent. Although we may be equally happy to acquire a $100 bill from a person who hands it to us and from someone who accidentally drops it from his wallet while running by, most of us would be inspired to cooperate with the first and not with the second. Tamarins see the world in the same way.

Tamarins give food to unrelated others. Giving is not the simple by-product of feeling good. Giving depends on whether food was given in the past, and how it was given. Although these findings address many of the essential ingredients of Trivers's account, a closer look at the patterns of giving reveal the signature of an unstable system. When the tamarins play against the unilateral altruist, they don't respond with unilateral cooperation as one might expect given the level of generosity. They cooperate less than 50 percent of the time, and as each game progresses, the amount of food giving drops. This decline represents the signature of most games of cooperation developed by economists. If we repeatedly cooperate to achieve some goal—hunting a gazelle, let's say—then it pays to defect on the last interaction because the relationship is ending. But if I think through this logic right before the last opportunity to interact, then I will surely think about defecting on the second to last opportunity, and the third to last, and so on. Cooperation unravels. We see this same unraveling with the tamarins. A further sign of vulnerability comes from the experiments with the untrained pairs. Here, cooperation ends if one of the players defects on two consecutive opportunities to pull, a pattern that happens often. Like the jays,

tamarins can maintain some level of cooperation under some restricted conditions. Overall, however, it is an unstable system. And although tamarins naturally give food to each other in the wild and in captivity, providing a certain level of ecological validity, tamarins in nature will rarely have the opportunity to reciprocate with unrelated animals, with the exception of their mates and possibly a sneaky mating with a neighbor. Thus, although tamarins have evolved some of the necessary psychological ingredients for reciprocal altruism—detecting cheaters, calculating contingencies, distinguishing between accidental and intentional actions—it is unlikely that these mechanisms evolved for reciprocal altruism. They may, however, have evolved to solve kin-based interactions.

How Common Is Reciprocity?

Despite these examples and the enormous theoretical interest in cooperation and reciprocity in the Prisoner's Dilemma, very little empirical evidence supports the theory for nonhuman animals. Most instances of putative reciprocity have either not been replicated or can be accounted for by simpler mechanisms. For example, one of the first observations of reciprocity involved alliance formation in olive baboons (Packer, 1977). Packer found that when soliciting help to separate a female from a mating pair, males tended to choose the same partners and alternate which partner solicited help to gain access to the female. A similar analysis on the closely related yellow baboon failed to find this reciprocal relationship (Bercovitch, 1988), and another study suggested that coalition formation is not a Prisoner's Dilemma at all (Noë, 1990). Another case of reciprocity involved predator inspection in fish (Dugatkin, 1991; Milinski, 1987). Milinski and Dugatkin have contended that some species of fish take turns potentially risking their safety by approaching predators. However, others suggest that this behavior can be explained by simpler mechanisms such as group cohesion (Lazarus & Metcalfe, 1990; Stephens, Anderson, & Benson, 1997). None of these point-counterpoint cases settles the issue. They do, however, leave us with doubts concerning the significance of reciprocity for animal social relationships.

Given the theoretical feasibility and the ubiquity of human reciprocity, why do we not find much evidence of nonhuman animals reciprocating (Hammerstein, 2003)? We propose two hypotheses: (1) researchers have not used the appropriate species and/or methodology to find reciprocity or (2) reciprocity is too cognitively demanding for most, if not all nonhuman animals. The first hypothesis is unlikely for a number of reasons. First, theory predicts that reciprocity could apply to a variety of species from bacteria to primates—theoretically, there are few constraints limiting which species should be able to reciprocate. Because of this, cooperation has been investigated in a large number of species including invertebrates, fish, birds, and many mammal species. Second, researchers have used a number of methodological tech-

niques to investigate cooperation, ranging from observation to natural experiments to highly controlled laboratory experiments. The breadth of species and techniques used suggests that the lack of evidence for reciprocity is not because we have not looked carefully.

We propose that reciprocity is a deceptively simple-sounding strategy that is intuitively appealing but cognitively complex. Reciprocity may, in fact, require a complicated suite of cognitive abilities that may limit its utilization by many animal species.

Cognitive Constraints on Cooperation

Cognitive abilities are clearly important in constraining animal behavior, and a resurgence of interest in integrating proximate and ultimate questions proves this point (Krebs & Davies, 1997). To investigate how these constraints influence cooperation and reciprocity, we must break this difficult problem down into its component parts. Here we examine a suite of cognitive abilities necessary to implement reciprocal strategies. Although reciprocity may tap cognitive abilities such as memory, cheater detection, social learning, and theory of mind, we focus on three constraints supported by a considerable amount of evidence: inhibitory control, temporal discounting, and numerical discrimination.

Inhibitory Control

Consider the following deal. In my left hand I am holding a $100 bill and in my right, a $1 bill. I say nothing at all, but give you an inviting wink that suggests you should reach for the money. Your instinct is surely to reach for the $100 bill. When you do, I pull this hand back and offer you the $1 bill. What would happen the second time around? Would you reach for the $100 bill again, or switch strategies and see what happens? Puzzled by the first outcome, you might reach again for the $100, and then rediscover that you are offered the $1 bill. Soon enough, you would switch, reaching for the $1 bill and obtaining the $100 bill. You would now have a hunch about what is going on. You would surmise that the best strategy is to pick the amount you don't want to get the amount you want. You have acquired a new rule to solve this task. But you have accomplished this task by overriding an old rule. We believe that this kind of problem, which requires resolving conflict and inhibitory control, is a core component of reciprocation, and may provide one explanation for why animals have difficulty.

The psychologist Sally Boysen (see chapter 10, this volume) ran this experiment with adult chimpanzees using food treats rather than dollar bills. Boysen assumed that chimpanzees were highly motivated for food and would choose the hand with less food to obtain the hand with more food. Precisely the opposite occurred. The choosers reached for the larger quantity of food and consistently received the smaller quantity.

We can explain this pattern of choice in two ways: either the choosers are exceptionally altruistic, intending to give away the larger stash of food, or they are incapable of controlling their desire to reach for the larger amount of food. The inhibitory challenge here is motivational, as the chimpanzees have evolved brains and stomachs designed to maximize the amount of food consumed. But Boysen's "reverse contingency" task requires the chimpanzees to overcome either an overlearned or innate response to reach for larger quantities of food over smaller quantities. If two simultaneously available patches of food differ only in terms of their quantities, no animal would feed in the patch with less food. But this is precisely what Boysen's task demands. Apparently, these chimpanzees are incapable of overriding the desire to reach for more food.

An alternative explanation is that Boysen's assumption that the chimpanzees were motivated for food was invalid. After all, they are fed every day, and even if they point to the larger quantity of food, they are still rewarded with the smaller quantity. Would they eventually learn to point to the smaller quantity of food if they received nothing at all after pointing to the larger quantity? Can other species solve this task? When other species—Japanese macaques, squirrel monkeys and cotton-top tamarins—are tested on Boysen's task, all fail, picking the larger quantity of food and getting stuck with the smaller quantity (Anderson, Awazu, & Kazuo, 2000; Kralik, Hauser, & Zimlicki, 2002; Silberberg & Fujita, 1996). But if the experimenter imposes a cost, withholding all food when subjects pick the larger quantity, macaques and squirrel monkeys eventually learn to pick the smaller quantity; tamarins stick with the losing strategy, picking the larger and getting nothing at all. This suggests that part of the chimpanzee's failure to point to the smaller quantity is due to the lack of costs associated with pointing to the larger quantity.

To solve Boysen's task, individuals must first inhibit the impulse to reach for the larger quantity and second, reach for the smaller but less desirable quantity. The primates' failure appears to be due to their underlying motivation for more food. This hypothesis is strengthened by a second set of experiments. Boysen ran a different version of the original task with chimpanzees that already knew the Arabic symbols from 1 to 6. This time, instead of choosing between one food treat versus four treats, the chimpanzees chose between a card with the number "1" written on its face and a card with the number "4"; each card covered up the corresponding number of food treats. The chimpanzees quickly learned to pick the number 1 card and received four treats, indicating chimpanzees can learn a rule like, "Point to the one you don't want to get the one you want." Therefore, it seems as though the difficulty of this task results from the chimpanzees' strong motivation to reach for food rather than an inability to learn the reverse-contingency rule.

Although Boysen's task does not pose a cooperative dilemma, it does set up an inhibitory problem that individuals must solve in order to stabilize cooperation. The

chooser must reach for the undesirable over the desirable food quantity, and then wait a return. The first move can be likened to giving away food, the second step to waiting for a reciprocated act (see the following discussion). The first move is costly to self, the second is beneficial. In Boysen's task, the cost appears too great, the inhibitory system too weak.

Temporal Discounting

A related topic to the inhibition problem is temporal discounting (also considered in studies of delayed gratification, impulsivity, and rate maximization)—a devaluing of future rewards. Discounting often results in a preference for smaller, immediate rewards over larger, delayed rewards. For example, imagine that a monkey encounters an unripe fruit. Should it consume the fruit now or wait for it to ripen (Kacelnik, 2003)? Waiting would yield a higher fitness benefit (more sugars are available), but the future is uncertain—another monkey may eat it; winds may knock it into a stream below; a fungus may infest it, spoiling a perfectly good fruit. This uncertainty may have provided a strong adaptive benefit for a preference for immediacy.

Given that the future is uncertain, should all organisms discount in the same way? Although impulsivity is probably universal among animals, the *rate of discounting*—that is, how quickly animals devalue food over time—varies widely across species, ages, and even context. In experiments that estimate discounting rate, subjects are presented with two stimuli: one associated with a small, immediate reward and the other with a large, delayed reward. The discounting rate is "titrated" by incrementally increasing the delay-to-large until subjects are indifferent between choosing the large delayed reward and the small, immediate reward. Therefore, researchers can find indifference points between immediate and delayed rewards over a range of small and large reward amounts. Pigeons and rats both discount future rewards quite highly (Mazur, 1987; Richards, Mitchell, de Wit, & Seiden, 1997)—sometimes devaluing a reward by up to 50 percent in the first second of delay!

Humans have a much lower discounting rate (Rachlin, Raineri, & Cross, 1991); it is by no means constant, however. In fact, impulsivity changes with age—children are much more impulsive than adults (Green, Myerson, & Ostaszewski, 1999). Mischel and colleagues have actually followed children to adulthood, measuring their impulsivity longitudinally (Mischel, Shoda, & Rodriguez, 1989). They have found that, although impulsivity decreases with age, impulsive children tend to develop into impulsive adults. In addition, impulsivity at a young age provides a reasonable predictor of future intelligence, social responsibility, resistance to temptation, and response to stress.

Discounting is not necessarily a static parameter that applies to any choice situation. Rather, it can change choice preferences in different situations, thus the discounting rate is context-dependent. For example, blue jays are usually quite impulsive

birds, preferring immediate to delayed rewards (Stephens & McLinn, 2003). In the autumn, however, jays switch from consuming every acorn they encounter to caching them behind tree bark or under leaf litter. This example of context-specific discounting is common across a number of bird and mammal species. The economic paradigm used to assess discounting also has profound effects on choice. Stephens and McLinn (2003) found that in three potentially equivalent economic scenarios,[2] blue jays acted impulsively in one and exhibited more self-control in the other two. The time from choice to reward greatly influenced discounting behavior in the jays even when the overall reward rate was the same. The extreme variation found in human discounting rates (Frederick, Loewenstein, & O'Donoghue, 2002) may also be attributed to our sensitivity to the economic context.

Many psychologists consider the iterated Prisoner's Dilemma as a type of discounting problem (Green, Price, & Hamburger, 1995; Rachlin, 2000). Individuals can choose between the immediate reward of defecting (and gaining only *P* fitness units) or the long-term reward of cooperating (and gaining *R* fitness units). Cross-sectional data on human discounting and cooperation agree with this perspective. After using a titration experiment to establish individual discounting rates, Harris and Madden (2002) found that these discounting rates correlated with cooperation levels—less impulsive individuals cooperated more frequently.

Although Axelrod and Hamilton (1981) included a discounting parameter in their original formulation of tit-for-tat, their parameter only considered the probability of future interactions. Experimental evidence indicates that sheer repetition is not enough to circumvent the discounting problem; other methods must be used to mitigate discounting. As mentioned earlier, Stephens and colleagues (2002) offered repeated Prisoner's Dilemma games to blue jays, using the same general approach as in the Clements and Stephens's studies (figure 9.2). This time, however, they altered the delay to payoff. They only found cooperation when the payoffs accumulated over several trials. This accumulation technique reduced impulsivity in a discounting game as well, suggesting that the jays could only cooperate when their natural discounting tendencies were reduced. Baker and Rachlin (2002) also reduced discounting in pigeons by decreasing the time between Prisoner Dilemma trials. When these intertrial intervals were short, the pigeons cooperated more frequently.

Recent theoretical and empirical evidence indicates that temporal discounting can have profound influences on cooperation. This interface between biology, economics, and psychology provides rich opportunities in which to ask important questions about the nature of social choices, including the mechanisms that both facilitate and constrain them. Beyond the theoretical insights we have sketched, these findings also imply that discounting must be considered when designing appropriate tests of cooperation.

Numerical Discrimination

A challenge for reciprocal interactions lies in quantifying the economics of the entities given and returned, and evaluating whether the exchange was equitable. For example, if an altruist gives four apples and receives one back, this is not equitable, and natural selection should eliminate this poor decision maker from the population. Do animals count or quantify in these ways? If they do, then individuals can at least assess equitable returns. If they don't, then either individuals are satisfied with some return, regardless of amount, or they are open to defectors' giving back less than a fair amount. The simple answer is that animals can count small numbers less than four with precision, and large numbers greater than four with only an approximate sense, with evidence for these two systems anchored in a wealth of behavioral and neurobiological studies (see chapters 6, 7, and 8).

Rats and pigeons can be trained to press a key for food. They can also be trained to press the same key a number of times for food. However, animals make more errors as the required number of presses increases (reviewed in Boysen & Capaldi, 1993). When an experimenter requires a rat to press a key four times for food, it is usually dead on, pressing exactly four most of the time, and on occasion, pressing three or five times. In contrast, when the target number is 30, sometimes the rat presses 30 times, but often it presses somewhere between 20 and 40 times. We observe these patterns when the rat (or pigeon) has to count the number of presses, light flashes, or tones, and when the task changes from counting to waiting a particular period of time before pressing. What these studies show is that animals can count, but only approximately so. In terms of reciprocal exchanges, animals can quantify, approximately, what was given and what was received. But does this map on to anything in the natural world?

There are at least three naturally occurring situations where number would appear to matter: aggressive competitions within and between groups, foraging for food, and reciprocal exchanges of resources in either the same or different currencies. In lions and chimpanzees, two species that attack and kill foreign intruders, individuals attend to the number of competitors. In lions, playback experiments in Tanzania show that females respond more aggressively to one foreigner calling than to three foreigners calling (McComb, Packer, & Pusey, 1994). In chimpanzees, a group of males is more likely to approach and kill a foreign male if the attackers have a three to one advantage (Wilson, Hauser, & Wrangham, 2001). Within groups of dolphins, lions, and many primate species, two or three individuals will form coalitions to defeat either a single dominant individual or a smaller coalition. Although these coalitions involve small numbers, they nonetheless require some capacity to count the number of competitors. And this capacity emerges in a naturally occurring, evolutionarily significant context.

Studies of foraging in animals show that individuals attempt to maximize the rate of energetic returns, picking patches with more over less food (Stephens & Krebs, 1986). Since estimates of rates of return depend on quantifying the amount of food consumed over time, we can ask whether animals count the pieces, guesstimate the volume, or time the foraging periods in a patch. Hauser and colleagues (2000) showed lone rhesus monkeys two opaque empty boxes and then put two pieces of apple into one box and one piece of apple into the other. Subjects consistently picked the box with two apples; they also picked the box with three apples over two, and four apples over three. But when they were presented with five versus four apple pieces, some animals picked four and some five. Without training, rhesus monkeys can count the number of pieces of food, and spontaneously discriminate four from three, but not larger numbers. But there's a problem. Perhaps the monkeys aren't counting at all. Perhaps they are timing how long it takes to load up the box with four apples versus three apples. Since it takes longer to place four pieces than three pieces, their internal timer will tell them to pick the box with four. Although timing is an impressive calculation, and one that plays an important role in the lives of all animals, it is not the same calculation as counting. But a simple experiment shows that number, not time, controls their response. If you place four pieces of apple into one box and three pieces of apple plus a rock into the other, equating time and the number of objects, rhesus pick the box with four pieces of apple. Number, not time, is responsible for the rhesus monkeys' preferences.

Based on an overwhelming number of carefully controlled experiments, it is now fair to say that animals have a number sense. It is a capacity that, in our opinion, consists of two naturally available systems (for a recent review of this literature, see Hauser & Spelke, in press; for a different perspective, see chapter 6, this volume). One allows animals to count up to about four with precision; the second allows them to approximate number, but without any limits on magnitude. Humans, including both young infants and mature adults, also have these two systems. But we have an additional system that relies on language, and in particular, words for numbers. This third system allows us to discriminate any two numbers with precision.

Returning to the problem of reciprocal altruism, we can now give a more specific answer to the question of limits or constraints on reciprocation, especially with respect to quantificational abilities. If and when animals engage in a bout of reciprocal altruism, they will either be limited to small numbers of objects in cases where the exchange must be precise (a banana for a banana), or they will be freed from this constraint where approximate exchanges are tolerated. The same prediction holds for cases where the currency is time, such as the duration of a grooming bout. If one antelope grooms another for 10 minutes, the groomer will most likely accept—as fair exchange, that is—a reciprocated grooming bout of between 8 and 12 minutes. A cheater who shoots for a slightly shorter bout would, in the long run, win. And

this selfish victory might well cause the demise of a potentially cooperative society.

Neural Correlates of Cooperation

Given the cognitive constraints discussed, implementing reciprocal cooperative strategies may prove difficult for many animal species. In cases in which we do see reciprocation, the question remains: are reciprocators integrating all of these domain-general abilities or are they tapping domain-specific cognitive adaptations for cooperation? In other words, are some animals specialized to reciprocate? One way to answer these questions is to delve into the brain to search for neural correlates of cooperative behavior. Although there are studies examining the neural correlates of the cognitive components such as inhibition (Hauser, 1999; Roberts & Wallis, 2000), temporal discounting (Manuck, Flory, Muldoon, & Ferrell, 2003), and numerical competence (Nieder, Freedman, & Miller, 2002; Sawamura, Shima, & Tanji, 2002), there are no studies exploring the neurobiology of cooperation in nonhuman animals. For this reason, here we focus on neuroeconomics—the neurobiology of economic decision making in humans—with the hope that it will shed some light on the possible neural correlates in animals, and minimally, open the door to research in this area. The logic is basically this: if we can document the necessary and sufficient circuitry underlying human cooperation and reciprocal interactions, then this provides one way in which one species solved the problem of reciprocal altruism. Although other animals may solve this problem by means of other circuitry, if animals are incapable of maintaining reciprocally stable relationships, then understanding which part of the circuitry is missing or deficient may help explain why.

Neuroeconomics of Cooperation

The emerging field of neuroeconomics integrates the latest technology of functional neuroimaging and neuronal recordings with classical experimental economics to determine how the brain makes economic decisions. Recently, economists and neuroscientists have collaborated to perform functional neuroimaging on human subjects playing several different cooperative games. In all of these games, individuals can choose to behave selfishly or altruistically. Because each game is either sequential or repeated, recipients of the altruistic option have the opportunity to reciprocate—therefore, these are potentially games of reciprocity. Because this is a nascent field, the data are limited to a few studies, and the causal relationship between brain area and function are still unclear.

Rilling and colleagues (2002) used functional magnetic resonance imaging (fMRI) to scan subjects that played repeated sequences of the Prisoner's Dilemma game shown in figure 9.4A. Subjects played against three partners: a freely behaving human, a

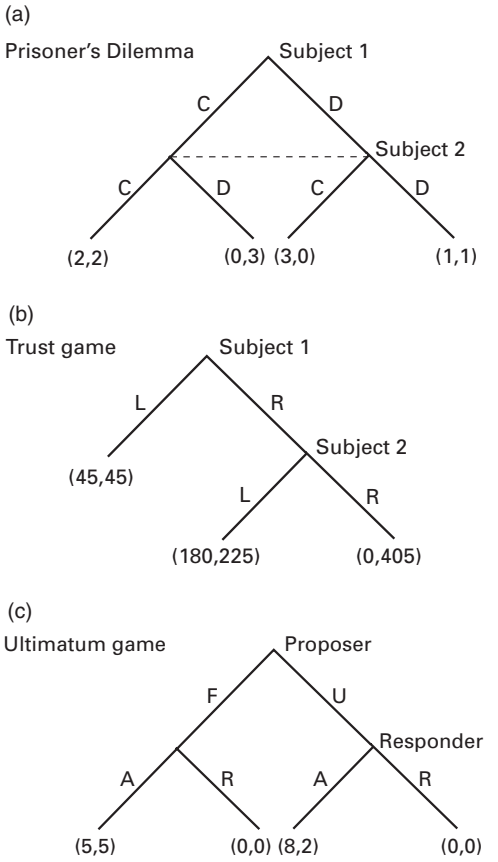


Figure 9.4

Extended form of games in neuroeconomics experiments depicting the sequential choices of the players. The nodes of the trees represent choice points for particular players. Dashed lines connecting nodes indicate that the player does not know her current position in the decision tree. Payoffs are indicated in parentheses for subject 1 and subject 2, respectively. (*A*) In the Prisoner's Dilemma, both subjects simultaneously choose either to cooperate (C) or defect (D). In this example, subject 1 chooses C or D, and subject 2 chooses not knowing subject 1's choice. The equilibrium strategy is mutual defection for a one-shot game. (*B*) In the trust game, subject 1 can choose left (L) and quit the game with a small payoff or choose right (R) to continue the game. Subject 2 can then reciprocate subject 1's trust by choosing L to receive a moderate payoff or break subject 1's trust by choosing R to receive a large payoff. The equilibrium strategy is for subject 1 to choose L and quit the game. (*C*) For the ultimatum game, the proposer submits either a fair (F) or unfair (U) offer. The responder can either accept (A) that offer and distribute the earnings or reject (R) the offer, discarding the earnings for both players. Proposing an unfair offer and accepting any nonzero offer are the equilibrium strategies. Iterated play can lead to stable cooperation and reciprocation in all three games.

human stooge that began cooperating but defected after three consecutive mutually cooperative moves, and a computer playing tit-for-tat. Against both human partners, subjects initially cooperated but reduced cooperation in later trials. When playing the computer, subjects initially defected, then increased cooperation, only to defect in the final trials. Mutual cooperation with human partners was associated with activation in both the anteroventral striatum and the orbitofrontal cortex (Brodmann Area 11) more than the other three possible outcomes; both of these areas have been implicated as playing a significant role in reward (Damasio, 1994; Rolls, 1999). Only the orbitofrontal cortex, however, was activated by mutual cooperation with a computer partner. Reciprocating a partner's previous cooperation increased activation in the anterior caudate, the postcentral gyrus (BA 1/3), the anterior cingulate cortex (BA 25), and the anteroventral striatum. Again, these areas play a role in reward assessment, and significantly, in resolving conflict when there are alternative moves. In our opinion, these data do not provide any additional explanatory power with respect to the psychology and economics of decision making. They do, however, provide new insights into the neural correlates, and these are important with respect to both comparative and developmental data: Are similar areas activated in animals playing the same game? What about human children at different ages?

In the Prisoner's Dilemma, both subjects choose simultaneously and therefore do not know their partner's choice. In a closely related "trust" game, subjects move sequentially, so that the second player to choose can reciprocate the first player's kindness. For example, McCabe and colleagues (2001) imaged subjects playing the trust game shown in figure 9.4B. Subjects played against a human partner and against a computer playing a known, probabilistic strategy. Although they do not describe the behavioral results, previous studies show that the first player cooperates in about 50 percent of the trials, and the second player reciprocates in 75 percent of those trials (McCabe & Smith, 2000). No data were presented on contrasting behavioral responses to human and computer partners. Subjects that cooperated on at least one third of the trials showed increased activation in the occipital lobe (BA 17, 18), the parietal lobe (BA 7), the thalamus, the middle frontal gyrus, and the frontal pole (BA 10) when playing against a human rather than a computer. The authors suggest that because of selective activation in the prefrontal cortex specifically when cooperating with a human partner, the trust game may recruit theory of mind modules such as shared-attention mechanisms. Although intriguing, these results leave several issues unresolved. For example, does the trust game require a theory of mind? Would autistics, who show a deficit in theory of mind type tasks, make similar decisions as do nonautistics in the trust game? What about normal children, below the age of approximately five years, who have yet to acquire a full-blown theory of mind? Given that the prefrontal cortex also plays a significant role in inhibition (see above), and that one of the problems facing both autistics and young children with respect to theory of mind

tasks is an inhibitory one, could prefrontal activation reflect inhibitory mechanisms as opposed to theory of mind systems? Why do the other areas of the brain show high levels of activation only against human partners?

A commonly analyzed cooperative game—the Ultimatum Game—examines how individuals value fairness. In this game, the first player (the proposer) is given an amount of money to split between herself and the second player (the responder). After proposing a split, the responder can either accept the offer or reject it, thereby preventing either player from receiving any money (figure 9.4C). Economic theory predicts that the proposer's offer should be the smallest possible amount to the responder, and the responder should accept any positive offer. After all, both players benefit—why look a gift horse in the mouth? Surprisingly, in experimental game situations, proposers tend to offer approximately 50 : 50 splits, and responders often reject offers lower than 20 percent of the stake (reviewed in Camerer, 2003).

Sanfey and colleagues (2003) scanned subjects playing ultimatum games against human and computer partners. The human and computer partners acted only as proposers, playing a fixed strategy of offering either fair 5 : 5 splits of ten dollars or offering unfair splits of 9 : 1, 8 : 2, or 7 : 3. Subjects acted as responders, accepting the fair offer and increasing rejection as the offers became more unfair; however, they treated human and computer proposers differently by rejecting more unfair offers from human partners. Unfair offers from human partners were associated with activation in the bilateral anterior insula, dorsolateral prefrontal cortex (BA 46), and anterior cingulate cortex (BA 24/32); again, these areas are significantly involved in emotional regulation, choice and in resolving conflict. There was greater activation in the anterior insula for unfair offers from human partners than computer partners. Insular activation was also associated with rejecting unfair offers; the authors contend that this activation corresponds to negative emotional states such as pain, distress, anger, and especially disgust. They also state that activation in the dorsolateral prefrontal cortex indicates a conflict between its "executive control" functions and the insular emotional reactions. At the level of function, these results are not surprising. Playing the Ultimatum Game involves resolving a conflict between selfishly keeping the larger proportion of the initial pot and being generous; it also involves emotion, especially on the part of the recipient who obtains either a fair or unfair offer. It was therefore expected that the circuitry underlying decision making, conflict resolution, and emotion would activate in this task. Nonetheless, this study pinpoints some of the necessary substrates for cooperation, opening the door to both comparative and developmental studies.

Cooperation and the Brain

Since neuroeconomics is in its infancy, there is no clear computational theory predicting how cooperation is processed and represented in the brain (but see McCabe, 2003). This is evident by the piecemeal findings of these first studies of the neuroe-

conomics of cooperation. In all of the games presented here, economic theory predicts selfishness, but experimental results show that people tend to cooperate well above expected levels. Despite these similarities in behavior, the neuroeconomic studies reviewed here all cite different cognitive components in their results: reward-center processing, executive control, emotional centers, and conflict-resolution areas. Perhaps these disparities simply reflect the growing pains associated with integrating neuroscience and economics, especially the psychology of decision making and choice (Glimcher, 2003).

Although neuroimaging studies can provide interesting correlates of behavior, causality is difficult to infer. Neuronal recordings of candidate brain centers in nonhumans can offer more direct assessment of neuronal activity and can provide an evolutionary framework for understanding the cooperative brain. One potentially fruitful avenue might be through the mirror neuron system located in the premotor cortex of macaques and humans (see chapter 11 for specific details). This part of the brain is equipped with neurons that fire when an individual performs a particular action or when the subject sees an individual perform the exact same action. These neurons fire for action and perception. We tentatively propose that they could provide a necessary, but not sufficient piece of circuitry for reciprocal altruism. For example, there are mirror neurons that only respond when a hand grasps a piece of food, rotates around a piece of food, displaces food, or releases food. Other neurons fire only when a hand grasps an object with index finger and thumb, and not at all when a pair of pliers, held by a hand, grasps the object in the same way. Some neurons even fire when the complete trajectory of the action is concealed, thereby causing the animal to infer or predict the intended action relative to the target object. Together, the fine coding of these mirror neurons suggests that the premotor cortex provides a warehouse of motor commands, a library of action manuals. Given the symmetry between action and perception in the tamarin task, it is possible that the mirror neuron system was engaged. Actor A pulls the tool and gives food to animal B. As animal B watches, A's pull triggers a matched response in B, thereby beginning the mirror neuron loop. This explanation might provide the most parsimonious explanation for the origins of this system. Others have argued that it evolved for imitation and theory of mind, two capacities that macaques lack. It would therefore be of interest to run the tamarin reciprocation experiment while individuals are in a scanner. The strong prediction is that seeing one's partner pull will activate the mirror neuron system as the perceiver's action system clicks into gear.

Conclusions

Cooperation is quite common in both human and nonhuman societies (Dugatkin, 1997). We argue, however, that most instances of animal cooperation can be attributed to selfish benefits or indirect benefits via helping kin. True altruistic cooperation

maintained by reciprocity is rare if not absent among animals, despite its ubiquity in humans. We propose that cognitive constraints on animal inhibition, temporal discounting, numerical discrimination, memory, cheater detection, punishment, theory of mind, and other components may limit the ability of many species to implement and maintain reciprocally altruistic strategies. In particular, animals have difficulty inhibiting the tendency to choose large amounts of food when available. This inhibition problem could pose a challenge for making altruistic decisions that require forgoing large rewards for smaller rewards. In addition, animals often highly discount future rewards. The extreme preference for immediacy exhibited by many species makes waiting for reciprocated rewards very difficult. Finally, precise numerical competence in animals is restricted to small quantities—larger quantities are estimated. When exchanging rewards in cooperative situations requires precision, the quantities in question may be limited to small numbers. These and other faculties may be necessary components of our capacity to reciprocate. If correct, then comparative research must illuminate which components are shared with other animals, which are unique to humans, and why certain components evolved in our species and no other. Although the crocodile, lion, chimpanzee, and Maasai warrior may all cooperate during a hunt, only the Maasai may engage a uniquely human, domain-specific specialization for cooperation.

We can now return to a question raised earlier on: What kinds of cognitive specializations, if any, are required for reciprocation? One approach to answer this question is to assess what happens in the brain when individuals are placed in cooperative games. Although we know little about how animals make decisions in these games, the burgeoning field of neuroeconomics is elucidating the role of the human brain in reciprocal games. Neuroimaging studies of the Prisoner's Dilemma, trust games, ultimatum games, and investment games implicate various areas of the prefrontal cortex, cingulate cortex, and striata among others. While interesting, these studies provide only correlational inferences concerning the relationship between brain activation and decision-making behavior. The timing of activation—whether it occurred at the time of the decision or is a consequence of the decision—is unclear. The real power of these studies lies in their implications for comparative and developmental questions. Does activation correlate with decision making in nonhuman animals playing similar games? Does activation change over the lifespan of an individual human? That is, as these brain centers come online in children, how does their decision making change? These types of questions get at the heart of a cognitive theory of reciprocity.

These cognitive and neurobiological analyses provide interesting insights into the economics, psychology, and evolution of altruistic cooperation. Ultimately, understanding the nature of human cooperation will require cooperation among disciplines.

Acknowledgments

This work was supported by an NIH-NRSA fellowship to JRS and an NSF-ROLE grant to MDH. We thank Stanislas Dehaene and Andreas Nieder for comments on the chapter.

Notes

1. A growing group of economists and anthropologists have argued that Triversian reciprocity is actually selfish as the initial altruistic act is made with the explicit expectation that the recipient will return the favor. Humans appear to have evolved a different, and apparently unique form of cooperation called *strong reciprocity*, defined recently by Gintis and colleagues (2003) as a "predisposition to cooperate with others and to punish those who violate the norms of cooperation, at personal cost, even when it is implausible to expect that these costs will be repaid either by others or at a later date."

2. By "potentially equivalent economic scenarios" we mean that the overall intake rate for the animals is the same for each scenario given a particular strategy. For example, if the animal chooses the small reward every time, it would receive the same amount of food over a given time period in all three scenarios.

References

Anderson, J. R., Awazu, S., & Kazuo, F. (2000). Can squirrel monkeys (*Saimiri sciureus*) learn self-control? A study using food array selection tests and reverse-reward contingency. *Journal of Experimental Psychology: Animal Behavior Processes*, *26*(1), 87–97.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*(4489), 1390–1396.

Baker, F., & Rachlin, H. (2002). Self-control by pigeons in the prisoner's dilemma. *Psychonomic Bulletin and Review*, *9*(3), 482–488.

Bercovitch, F. (1988). Coalitions, cooperation and reproductive tactics among adult male baboons. *Animal Behaviour*, *36*(4), 1198–1209.

Boysen, S. T., & Capaldi, E. J. (1993). *The Development of Numerical Competence: Animal and Human Models*. Hillsdale, NJ: Erlbaum.

Camerer, C. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

Clements, K. C., & Stephens, D. W. (1995). Testing models of non-kin cooperation: Mutualism and the Prisoner's Dilemma. *Animal Behaviour*, *50*(2), 527–535.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (pp. 163–228). Oxford: Oxford University Press.

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason and the Human Brain*. New York: Putnam.

De Waal, F. B. M. (1997). Food transfers through mesh in brown capuchins. *Journal of Comparative Psychology*, *111*(4), 370–378.

De Waal, F. B. M., & Berger, M. L. (2000). Payment for labour in monkeys. *Nature*, *404*(6778), 563.

Denault, L. K., & McFarlane, D. A. (1995). Reciprocal altruism between male vampire bats, *Desmodus rotundus*. *Animal Behaviour*, *49*(3), 855–856.

Dugatkin, L. A. (1991). Dynamics of the Tit for Tat strategy during predator inspection in the guppy *Poecilia reticulata*. *Behavioral Ecology and Sociobiology*, *29*(2), 127–132.

Dugatkin, L. A. (1997). *Cooperation Among Animals: An Evolutionary Perspective*. New York: Oxford University Press.

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, *425*(6960), 785–791.

Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review, 90*(4), 980–994.

Fehr, E., & Henrich, J. (2003). Is strong reciprocity a maladaptation? On the evolutionary foundations of human altruism. In P. Hammerstein (Ed.), *The Genetic and Cultural Evolution of Cooperation* (pp. 55–82). Cambridge, MA: MIT Press.

Feistner, A. T. C., & Price, E. C. (1990). Food-sharing in cotton-top tamarins (*Saguinus oedipus*). *Folia Primatologica*, *54*(1–2), 34–45.

Flombaum, J. I., Junge, J. A., & Hauser, M. D. (in press). Rhesus monkeys (*Macaca mulatta*) spontaneously compute addition operations over large numbers. *Cognition*.

Flood, M. M. (1958). Some experimental games. *Management Science*, *5*(1), 5–26.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *XL*(2), 351–401.

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior*, *24*(3), 153–172.

Glimcher, P. W. (2003). *Decisions, Uncertainty, and the Brain*. Cambridge, MA: MIT Press.

Green, L., Myerson, J., & Ostaszewski, P. (1999). Discounting of delayed rewards across the life span: Age differences in individual discounting functions. *Behavioural Processes*, *46*(1), 89–96.

Green, L., Price, P. C., & Hamburger, M. E. (1995). Prisoner's dilemma and the pigeon: Control by immediate consequences. *Journal of the Experimental Analysis of Behavior*, *64*(1), 1–17.

Hammerstein, P. (2003). Why is reciprocity so rare in social animals? A protestant appeal. In P. Hammerstein (Ed.), *Genetic and Cultural Evolution of Cooperation* (pp. 83–94). Cambridge, MA: MIT Press.

Harbaugh, W. T., Krause, K., & Liday, S. G. (2002). Children's bargaining behavior: differences by age, gender, and height. Working paper. Department of Economics, University of Oregon.

Harbaugh, W. T., Krause, K., Liday, S. G., & Vesterlund, L. (2003). Trust in children. In E. Ostrom & J. Walker (Eds.), *Trust and Reciprocity: Interdisciplinary Lessons from Experimental Research* (pp. 302–322). New York: Russell Sage Foundation.

Harris, A. C., & Madden, G. J. (2002). Delay discounting and performance on the prisoner's dilemma game. *Psychological Record*, *52*(4), 429–440.

Hart, B. L., & Hart, L. A. (1992). Reciprocal allogrooming in impala, *Aepyceros melampus*. *Animal Behaviour*, *44*(6), 1073–1083.

Hauser, M. D. (1999). Perseveration, inhibition and the prefrontal cortex: A new look. *Current Opinion in Neurobiology*, *9*(2), 214–222.

Hauser, M. D. (2000). *Wild Minds: What Animals Really Think*. New York: Henry Holt.

Hauser, M. D., Carey, S., & Hauser, L. B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proceedings of the Royal Society of London, Series B*, *267*(1445), 829–833.

Hauser, M. D., Chen, M. K., Chen, F., & Chuang, E. (2003). Give unto others: Genetically unrelated cotton-top tamarin monkeys preferentially give food to those who altruistically give food back. *Proceedings of the Royal Society of London, Series B*, *270*(1531), 2363–2370.

Hauser, M. D., & Spelke, E. S. (in press). Evolutionary and developmental foundations of human knowledge: A case study of mathematics. In M. Gazzaniga (Ed.), *The Cognitive Neurosciences III*. Cambridge, MA: MIT Press.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., et al. (2001). In search of Homo economicus: Behavioral experiments in 15 small-scale societies. *American Economic Review*, *91*(2), 73–78.

Kacelnik, A. (2003). The evolution of patience. In G. Loewenstein, D. Read, & R. Baumeister (Eds.), *Time and Decision: Economics and Psychological Perspectives on Intertemporal Choice* (pp. 115–138). New York: Russell Sage Foundation.

Kralik, J. D., Hauser, M. D., & Zimlicki, R. (2002). The relationship between problem solving and inhibitory control: Cotton-top tamarin (*Saguinus oedipus*) performance on a reversed contingency task. *Journal of Comparative Psychology*, *116*(1), 39–50.

Krebs, J. R., & Davies, N. B. (1997). *Behavioural Ecology: An Evolutionary Approach* (4th ed.). Oxford: Blackwell Science.

Lazarus, J., & Metcalfe, N. B. (1990). Tit for Tat cooperation in sticklebacks: A critique of Milinski. *Animal Behaviour*, *39*(5), 987–989.

Manuck, S. B., Flory, J. D., Muldoon, M. F., & Ferrell, R. E. (2003). A neurobiology of intertemporal choice. In G. Loewenstein, D. Read, & R. Baumeister (Eds.), *Time and Decision: Economics and Psychological Perspectives on Intertemporal Choice* (pp. 139–172). New York: Russell Sage Foundation.

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative Analyses of Behavior: The Effect of Delay and of Intervening Events on Reinforcement Value* (vol. 5, pp. 55–73). Hillsdale, NJ: Erlbaum.

McCabe, K. A. (2003). A cognitive theory of reciprocal exchange. In E. Ostrom & J. Walker (Eds.), *Trust and Reciprocity: Interdisciplinary Lessons from Experimental Research* (pp. 147–169). New York: Russell Sage Foundation.

McCabe, K. A., Houser, D., Ryan, L., Smith, V. L., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences*, *98*(20), 11832–11835.

McCabe, K. A., & Smith, V. L. (2000). A comparison of naive and sophisticated subject behavior with game theoretic predictions. *Proceedings of the National Academy of Sciences*, *97*(7), 3777–3781.

McComb, K., Packer, C., & Pusey, A. E. (1994). Roaring and numerical assessment in contests between groups of female lions, *Panthera leo*. *Animal Behaviour*, *47*(2), 379–387.

Milinski, M. (1987). Tit for Tat in sticklebacks and the evolution of cooperation. *Nature, 325*(6103), 433–435.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science, 244*(4907), 933–938.

Muroyama, Y. (1991). Mutual reciprocity of grooming in female Japanese macaques, *Macaca fuscata*. *Behaviour*, *119*(3–4), 161–170.

Nieder, A., Freedman, D. J., & Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science*, *297*(5587), 1708–1711.

Noë, R. (1990). A veto game played by baboons: A challenge to the use of the Prisoner's Dilemma as a paradigm for reciprocity and cooperation. *Animal Behaviour*, *39*(1), 79–90.

Packer, C. (1977). Reciprocal altruism in *Papio anubis*. *Nature*, *265*(5593), 441–443.

Rachlin, H. (2000). *The Science of Self-Control*. Cambridge, MA: Harvard University Press.

Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior*, *55*(2), 233–244.

Rapoport, A., & Chammah, A. N. (1965). *Prisoner's Dilemma: A Study in Conflict and Cooperation*. Ann Arbor: University of Michigan Press.

Richards, J. B., Mitchell, S. H., de Wit, H., & Seiden, L. S. (1997). Determination of discount functions in rats with an adjusting-amount procedure. *Journal of the Experimental Analysis of Behavior*, *67*(3), 353–366.

Rilling, J., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., & Kilts, C. (2002). A neural basis for social cooperation. *Neuron*, *35*(2), 395–405.

Roberts, A. C., & Wallis, J. D. (2000). Inhibitory control and affective processing in the prefrontal cortex: Neuropsychological studies in the common marmoset. *Cerebral Cortex, 10*(3), 252–262.

Rolls, E. T. (1999). *The Brain and Emotion*. Oxford: Oxford University Press.

Roush, R. S., & Snowdon, C. T. (2001). Food transfer and development of feeding behavior and food-associated vocalizations in cotton-top tamarins. *Ethology*, *107*(5), 415–429.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, *300*(5626), 1755–1758.

Sawamura, H., Shima, K., & Tanji, J. (2002). Numerical representation for action in the parietal cortex of the monkey. *Nature*, *415*(6874), 918–922.

Seyfarth, R. M., & Cheney, D. L. (1984). Grooming, alliances and reciprocal altruism in vervet monkeys. *Nature*, *308*(5959), 541–543.

Silberberg, A., & Fujita, K. (1996). Pointing at smaller food amounts in an analogue of Boysen and Berntson's (1995) procedure. *Journal of the Experimental Analysis of Behavior*, *66*(1), 143–147.

Stephens, D. W., Anderson, J. P., & Benson, K. E. (1997). On the spurious occurrence of Tit for Tat in pairs of predator-approaching fish. *Animal Behaviour*, *53*(1), 113–131.

Stephens, D. W., & Krebs, J. R. (1986). *Foraging Theory*. Princeton, NJ: Princeton University Press.

Stephens, D. W., & McLinn, C. M. (2003). Choice and context: Testing a simple short-term choice rule. *Animal Behaviour*, *66*(1), 59–70.

Stephens, D. W., McLinn, C. M., & Stevens, J. R. (2002). Discounting and reciprocity in an Iterated Prisoner's Dilemma. *Science*, *298*(5601), 2216–2218.

Stevens, J. R., & Stephens, D. W. (2004). The economic basis of cooperation: trade-offs between selfishness and generosity. *Behavioral Ecology*, *15*, 255–261.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*(1), 35–57.

Wilkinson, G. S. (1984). Reciprocal food sharing in the vampire bat. *Nature*, *308*(5955), 181–184.

Wilson, M. L., Hauser, M. D., & Wrangham, R. W. (2001). Does participation in intergroup conflict depend on numerical assessment, range location or rank for wild chimpanzees? *Animal Behaviour*, *61*(6), 1203–1216.

10 Do Monkeys Understand Actions and Minds of Others? Studies of Single Cells and Eye Movements

Erica N. Lorincz, Tjeerd Jellema, Juan-Carlos Gómez, Nick Barraclough, Dengke Xiao, and David I. Perrett

In order to understand the nature of visual processing that might support comprehension of actions, it is useful to first review the general scheme of cortical visual processing. Action coding appears to be prevalent within one area of the higher-level visual cortex labeled the STS (superior temporal sulcus). We also review the spectrum of cell types found in this area and the cell properties that apply to comprehension of the behavior of others. Finally, we consider the question of how the neural apparatus might translate into observable behavior guided by the actions of others.

Different Streams for Processing the Visual Stimulus

Existing ideas about where in the brain the processing of the different features of a visual stimulus occur are heavily influenced by the Ungerleider and Mishkin model (1982), and by a subsequent adaptation by Milner and Goodale (1995). The Ungerleider and Mishkin model envisages a separation of visual processing into two distinct cortical streams: a dorsal "where" stream, extending from V1 into the inferior parietal cortex, primarily dealing with the spatial relationships of objects, and a ventral "what" stream, extending from V1 into the inferior temporal cortex (IT) dealing with the shape and identity of objects. The role of the ventral stream in the recognition of complex objects is supported by findings showing a gradual increase in the complexity of stimuli analyzed by cells from V1 up to IT (Perrett & Oram, 1993).

Milner and Goodale questioned the strict "what-where" dichotomy, and suggested that space and form are processed in both parietal and temporal areas but for different purposes (Milner & Goodale, 1995). In their view, the ventral stream subserves visual "perception," that is, object and scene recognition, requiring allocentric spatial coding to represent the enduring characteristics of objects, while the dorsal stream subserves the visual control of "action," requiring egocentric spatial coding for short-lived representations (vision for perception versus vision for action).

These ideas of joint processing of form and position have been substantiated by studies at the cellular level, with cells coding for object shape in parietal cortex

(Murata, Gallese, Kaseda, & Sakata, 1996; Sereno & Maunsell, 1998), and cells coding for the object's spatial position in area V4 within the ventral stream (Dobbins, Jeo, Fiser, & Allman, 1998).

The Anterior Part of the Superior Temporal Sulcus (STSa) in the Monkey

The anterior part of the superior temporal sulcus (STSa) in the macaque monkey, corresponding to area STPa (Bruce, Desimone, & Gross, 1981), is considered part of the ventral visual stream. STSa cells often maximally respond to the visual appearance of the face and body and to body actions, most notably of conspecifics and humans (Földiák, Xiao, Keysers, Edwards, & Perrett, 2004). Gross et al. (1972) made the first startling finding of cells that responded selectively to the sight of a specific body part, for example, a monkey's paw. Subsequent work in STSa revealed populations of cells selective for the sight of faces and bodily actions, such as articulations of the limbs and torso, but also whole body actions such as walking (Bruce, Desimone, & Gross, 1981; Perrett, Rolls, & Caan, 1982; Perrett et al., 1984, 1985a, 1985b, 1989, 1992; Desimone, Albright, Gross, & Bruce, 1984; Jellema & Perrett, 2002; Jellema, Baker, Wicker, & Perrett, 2002).

Other STSa cells are tuned to multiple views of the same animate object or the same action (Logothetis, Pauls, & Poggio, 1995; Perrett et al., 1989; Jellema et al., 2002), or are tuned to conceptually related visual stimuli, such as multiple body signals of directed attention (Perrett et al., 1985a, 1992). This selectivity is most likely obtained through pooling of the outputs of cells coding for distinct stimuli related by temporal association.

Characteristic for many STSa cells is that they integrate information about form and motion of animate objects (Oram & Perrett, 1994, 1996; Tanaka, Koyama, & Mikami, 1999) and, as has only recently become clear, integrate information about the spatial location of animate objects (Baker, Keysers, Jellema, Wicker, & Perrett, 2000; Jellema & Perrett, 2003c). STSa cells often generalize their sensitivity to complex shapes across changes in various other stimulus properties such as size, retinal position, orientation, the species (human or monkey), luminance, and color (e.g., Perrett et al., 1984, 1989; Rolls & Baylis, 1986; Ashbridge, Perrett, Oram, & Jellema, 2000; Földiák et al., in press).

Collectively these summarized response characteristics suggest a role in object recognition and allocate the STSa to the ventral visual stream (Milner & Goodale, 1995). The findings in the macaque monkey have led to the idea that the STSa is primarily involved in the visual analysis of actions performed by other individuals (Perrett et al., 1989). This view is supported by recent brain imaging studies that show activation of the human posterior STS, which area is thought to be the homolog of the monkey STSa, for the perception of biologically significant stimuli. Examples of such

stimuli are "biological motion" of human figures, hand actions, static faces, eye gaze and eye motion, and meaningful actions (for reviews see Allison, Puce, & McCarthy, 2000; Puce & Perrett, 2003). A recent study explicitly showed that the human posterior STS represents goal-directed or intentional actions (Saxe, Xiao, Kovacs, Perrett, & Kanwisher, 2004).

Action Coding in the STS

Orientation of Body Cues

The visual information arising from body cues appears to contribute to STS cell sensitivity in a way that is consistent with the cell's role in analyzing the direction of attention. For example, cells tuned to the left profile view of the head are often additionally tuned to the left profile view of the body (Wachsmuth, Oram, & Perrett, 1994). However, the direction of another individual's head or body may not always be a reliable index of where that individual's attention lies (Perrett et al., 1992). Gaze direction seems a more powerful guide in this respect, and gaze may therefore also be expected to affect STSa cell responses. Indeed, tuning to both head view and gaze direction seems relatively common in the STS. Moreover, when conjoint sensitivity is observed, the effective gaze directions match the effective head directions, and gaze in an ineffective direction can prevent the response to an otherwise effective head posture. Thus cellular coding of head and gaze direction seems compatible with gaze direction taking precedent (Perrett et al., 1985, 1992).

Despite the findings of cellular sensitivity to attention direction in macaques, it still remains a matter of debate whether Old World monkeys are able to use information about the gaze direction of others. For example, Anderson, Montant, and Schmitt (1996) report that macaque monkeys cannot be trained to use human gaze to locate hidden food. Behavioral assessments in our lab, however, have shown that macaques do spontaneously utilize the direction of attention of conspecifics to orient their own attention (Emery, Lorincz, Perrett, Oram, & Baker, 1997; Lorincz, Baker, & Perrett, 1999). We will return to the issue of this discrepancy later in the chapter.

Modulation of Action Coding

Modulation by attention Cells in STSa that are selectively responsive to articulations of limbs or parts of the face or body, such as the mouth, eyes, head, torso, legs, arms, hands and fingers, have also been documented (Perrett et al., 1985b, 1990; Mistlin & Perrett, 1990). Cells may respond selectively to arm movements and not to equivalent leg movements, or to leg movements and not to equivalent arm movements (Jellema, Baker, Wicker, & Perrett, 2000). Additionally, the cells often showed sensitivity for the direction of the motion; some are tuned to reaching toward the observer,

others to reaching to the observer's left, and so forth. Movements directed away from the subject may acquire particular meaning in a given context. For example, if a food tray is kept out of sight to the subject's right, then the sight of the experimenter reaching right can become salient, since this may bring the experimenter's hand to food, which is subsequently given to the subject.

For a subset of the population of STSa cells responding to the sight of arm reaching, the response could be modulated by the direction of attention of the agent performing the action (Jellema et al., 2000). Actions performed when the agent faced and gazed in the direction of reaching were more effective than actions performed with the head and gaze oriented away from the direction of reaching. Body posture, which provides another potential cue to the direction of attention (Perrett et al., 1992; Wachsmuth, Oram, & Perrett, 1994), was found to contribute further to the "modulation" of the response to a reaching movement (Jellema et al., 2000).

Cells' responses with conjoint selectivity for congruent reaching and attention can be formed by combining the appropriate outputs of cells that respond to directed attention with the outputs of cells that respond to directed limb movements. The significance of someone's reaching toward an object while his or her attention is focused on the object clearly differs from an identical arm and hand action performed with attention directed elsewhere. In the former case, one is likely to infer that it was this person's intention to reach out for the object in order to pick it up or make contact with it. In the latter case, one may infer that the object was incidental to the arm extension. Crucial is that information about the reaching action, such as its direction and the possible presence of a reaching goal, is linked to information about the direction of attention of the performer. The responses of the cells to both the hand actions and to attention direction can be selective for movements of the agent that appear intentional. Actions that are attended to are more likely to be intentional, whereas actions accompanied by attention elsewhere can result in accidental effects.

Modulation by goals Some STSa cells seem to be sensitive to the causal relationship between an action and the object or goal of that action. This has been demonstrated most clearly in cells sensitive to purposeful hand-object interactions, such as reaching for, picking, tearing and manipulating objects (Perrett et al., 1989, 1990; Jellema et al., 2000). These STSa cells are sensitive to the form of the hand performing the action, and are unresponsive to the sight of tools manipulating objects in the same manner as hands. Furthermore, the cells code the spatiotemporal interaction between the agent performing the action and the object of the action. For example, cells tuned to hands manipulating an object cease to respond if (1) the object is removed, (2) the hand action is made in a direction away from the object, or (3) the hands and object move appropriately but remain spatially separated (Perrett et al., 1989). This selectivity ensures the cells are more responsive in situations where the agent's motion is *causally* related to the object's motion.

Modulation by location The brain integrates different features of a visual stimulus, such as its form, color, motion, and location, into a single coherent percept. Milner and Goodale suggested that space and form might be processed in both dorsal (parietal) and ventral (temporal) streams, but for different purposes (e.g., Goodale et al., 1991; Milner & Goodale, 1995). Recently we discovered cell populations in STSa that are sensitive to the spatial location of animate objects after they moved out of sight behind a screen (Baker et al., 2000, 2001).

More generally we found that spatial information is indeed integrated with form and motion information at the single cell level quite extensively within the STS. This capacity may enable STSa cell populations to form representations of goal-directed or socially relevant actions.

Jellema et al. (2004) tested cells responsive to the sight of walking (i.e., tuned to body view and direction of motion; Perrett et al., 1984) for sensitivity to the position of the walking person within the testing room. More than half of the cells were sensitive for the spatial location of the agent. Some cells would respond only to walking at a far distance, others only to walking nearby. The locations of the walking agent can be referenced to the subject's perspective: near or far from the subject. This assumes an egocentric frame of reference, but in principle the cells could just as well have used an allocentric frame of reference (i.e., spatial descriptions based on environmental landmarks rather than the subject's own position and orientation). Allocentric coding has been observed for STSa cells sensitive to goal-directed actions (Perrett et al., 1989), and for an STSa cell coding for occluded agents (Baker et al., 2001). Such relative positions are especially relevant in social interactions.

Spatial coding may indeed be widespread in STSa. Previously it was suggested that STSa plays a role not only in animate object identification but also in the visual analysis of the intentions and goals of others' actions, which forms an important aspect of social cognition (Emery & Perrett, 1997; Jellema & Perrett, 2002). The significance of spatial coding in STSa must be seen in this light. The spatial positions that individuals occupy with respect to each other, or with respect to objects, contain essential information for an observer when it comes to determining the goals and intentions of those individuals.

Temporal Associations in Action Coding

Implying next or past motion from current posture Actions performed by most animals typically involve articulation. To understand an articulated action performed by another individual, we do not necessarily have to witness the entire action sequence. A single momentary view is often enough to identify the likely action about to be performed, or recently executed. The same momentary view may permit the identification of the probable goal of the action. This is a very useful capacity since it

allows us to understand an agent's actions even when we get only one glimpse of an agent or goal before the agent becomes hidden from view.

Articulated motion seems to be preferentially processed in STS, as shown by Beauchamp, Lee, Haxby, and Martin (2003), who made a direct comparison between articulated and nonarticulated human motion, and found that the former activated the STS significantly more than the latter. Jellema and Perrett (2003b) studied the sensitivity of STSa cells to body postures containing implied motion and to actual articulated body movements. It was postulated that if the static articulated posture were to be presented in isolation (i.e., in absence of actual movement), STS cells would respond as if the associated actual motion was presented.

Articulated actions were defined as actions where one body part (e.g., a limb or head) moves with respect to the body part it is attached to; conversely, nonarticulated actions are actions where the equivalent body parts do not move with respect to each other but move as one. Similarly, articulated static body postures contain a torsion or rotation between parts, while nonarticulated postures do not (i.e., the head, chest and pelvis aligned and oriented in the same direction, typical of an "at rest" posture).

Testing cells that were responsive during particular actions revealed that 55 percent of cells responded both to the articulated action and to the articulated static posture that formed the endpoint of the action presented in isolation. The cells did not respond to the sight of the nonarticulated static posture, which formed the starting point of the action. Moreover the cells did not respond to static postures resembling the articulated end point posture, but that were in a more relaxed muscular state (i.e., nonarticulated). The cells also did not respond to other articulated body actions that were less often associated with the effective static articulated posture.

The above findings give rise to the intriguing possibility that the STSa cells code for a particular articulated action both when actually presented and when implied in a still image. Previously, STSa cell responses were described that were tuned to the same perspective view of multiple parts of the body (e.g., left profile view of the trunk and left profile view of the head; Wachsmuth, Oram, & Perrett, 1994). The cells described here required different perspective views of body segments (i.e., torsion or twisting of body parts). We suggest that the cells code for the implied motion contained in the static articulated posture, or, in other words, code for the association of motion and posture, rather than for the articulated posture per se.

The STS could support recognition of an object and the likely type of movement associated with that object. Representation of such an association could allow inferences as to whether or not the object was likely to be moving when it was briefly seen (or at the time a picture was taken), and its most likely previous and future trajectory.

The neural representations in STSa for *actual* biological motion may also extend to biological motion *implied* from static postures. The data show that the visual process-

ing of static form may contribute to the comprehension of dynamic actions. Sensitivity to associations between image form and motion could form the basis of the ability of the nervous system to retrieve likely motion implied entirely by static images.

Sequences of actions and postures Under natural viewing conditions, STSa cell responses to the sight of static body postures may be controlled by actions performed by that body in the one or two seconds immediately preceding the onset of the static posture (Jellema & Perrett, 2003a). In other words, the perceptual history can enable or prevent a cell's response to the current retinal input. For example, a cell may respond vigorously to the sight of, say, a face when the face was preceded by action A, but fail to respond to the identical face when preceded by action B and fail to respond to the sight of a face when presented without any preceding action. For 54 percent of cells sensitive to the static posture, the nature of the movement preceding a static posture proved critical to the modulation of cell responses.

These results show that the "vocabulary or grammar" of actions and body postures coded by single STSa cells is much larger than previously thought. Thus cells in temporal cortex could support the formation of expectations about impending behavior of others, which suggests a role in the understanding of actions. The neural representations for sequences of events may play a role in predicting or anticipating the next move or posture of the animate object. For example, the sight of a body that has just stopped walking forward may invoke an expectation that, should walking commence again, it is likely to resume a forward direction. The same view of a static body that has just stopped walking backward, by contrast, may be expected to move again in a backward direction should walking resume.

Action sequences that become hidden The actions of others are not always fully visible, for example someone may become hidden from our sight as they move behind a tree, or their hands may not remain fully in view as they reach to retrieve an object. Within STS it is now apparent that specific cell populations are activated when the presence of a hidden agent can be inferred from the preceding visual events (i.e., the agent was witnessed passing out of sight behind a screen and has not yet been witnessed reemerging into sight; therefore, the agent is likely to remain behind the screen; see Baker et al., 2001). The population response of STSa cells to this sequence of events is shown in figure 10.1. It shows that STSa cells responded maximally when individuals were seen to "hide" behind an occluding screen. In the three seconds following disappearance from sight behind a screen, the population response was larger than in the prior three seconds when the agent was visible and moving toward the screen. Some cells had no detectable response to visible movements but started responding 1–4 seconds after the agent had become completely hidden.

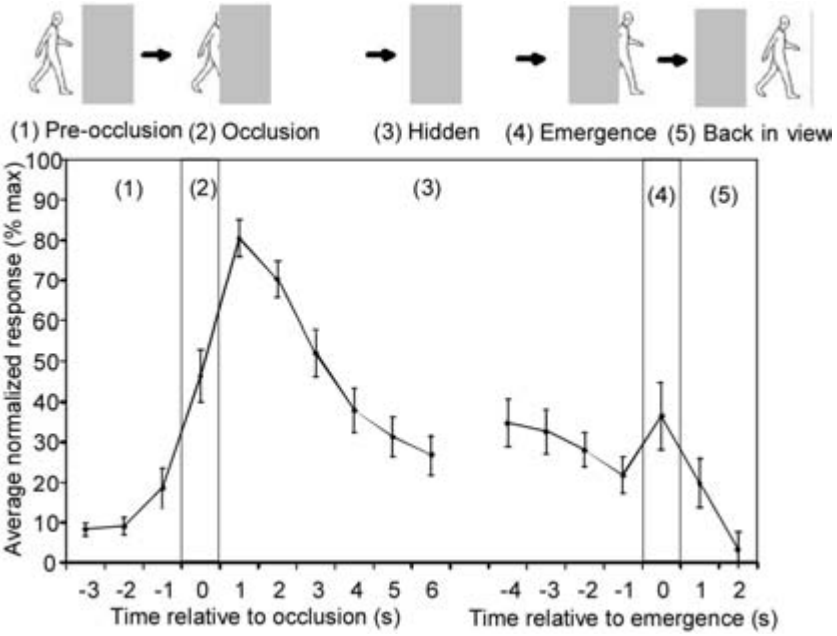


Figure 10.1
STS population response before, during and after an occlusion from sight. (*Upper*) Periods of the visual stimulus. The experimenter moved towards the occluding screen (1), was gradually occluded (2), remained hidden from view with only the screen visible (3), and gradually reemerged (4), until the experimenter was once again fully in view (5). Filled arrows show the progression of events. (*Lower*) Activity profile during the disappearance and subsequent emergence of the experimenter. The graph shows the average normalized population response of 26 cells recorded in STSa. On the left of the graph, responses are aligned with respect to the occlusion period. On the right, the responses are aligned with respect to the emergence period. (Adapted from Baker et al., 2001.)

The cells responding to occlusion additionally showed spatial sensitivity discriminating between locations at which the agent was completely hidden (at the left, right or middle of the room; Baker et al., 2001). Cell responses to the experimenter walking in-sight were consistent with the out-of-sight responses. For example, if hiding behind a screen located at the right-hand side of the testing room evoked significantly larger responses than hiding behind a screen at the left-hand side, then walking toward the right-hand screen would also evoke a larger response than walking toward the left-hand screen, with walking in both cases from left to right. These responses are consistent with the idea that the cells coded not only for the presence of the experimenter behind the right-hand screen, but also for the intention of the experimenter to go

behind that screen. For this interpretation, we need only assume that walking toward the right screen reflects the intention to move behind that screen.

The directional selectivity and position coding could be used to anticipate the reemergence of a hidden agent. Thus a monkey watching a human disappear behind a screen is aware in some sense that the human is hidden from sight and that he or she is likely to reappear from one or other side of the screen.

Relation of Visual Coding in STSa to Motor Planning in Premotor Cortex

The STSa cell populations coding body and hand actions appear to be predominantly sensory, although information from the motor system does affect other STSa cell populations (Hietanen & Perrett, 1996) and modulates STS activity in humans (Iacoboni et al., 2001; Nishitani & Hari, 2001). Gallese and Goldman (1998) suggested that the "action detecting" system in STSa could provide an initial "pictorial" description of the action, and this information is then relayed to frontal motor planning systems. The manner in which temporal STS and frontal systems interconnect is not fully clear, but probably involves intermediate processing steps mediated by parietal areas (Nishitani & Hari, 2000, 2001; Gallese, Fediga, Fogassi, & Rizzolatti, 2002).

The frontal region of primate cortex (inferior area 6) has long been known to be somatotopically organized for the representation and control of movements of the mouth and arm (Rizzolatti et al., 1988). This area can be subdivided into areas F4 and F5 (Rizzolatti & Gentilucci, 1988). Neurons in F5 are activated during specific motor acts performed with the hand or mouth, such as grasping, holding and tearing. Cells with responses related to reaching movements of the arm are typically found in F4.

Cells in F5 discharge during the execution of a particular action and during the sight of the same action. For example, an F5 cell, which responds selectively when the monkey executes a grasping action, may also respond (like STS cells) to the sight of another monkey or the experimenter grasping an object, but not to the sight of different hand actions such as tearing (di Pellegrino et al., 1992; Gallese et al., 1996; Rizzolatti et al., 1996a,b). An F5 cell selective for the execution of grasping would also respond when the monkey grasps an object in the dark, thereby demonstrating the motor properties of the response. These conjoint properties have led Rizzolatti et al. (1996a, 1996b) and Gallese et al. (1996) to postulate that the F5 neurons form a system for matching observation and executing actions for the grasping, manipulation and placement of objects. These neurons have now been labelled "mirror" neurons. Cells with response selectivity similar to those in F5 have recently been reported in the inferior parietal lobule (Gallese et al., 2002; see Williams, Whiten, Suddendorf, & Perrett, 2001).

The experiments in which actions are partially, or totally, occluded from sight have also highlighted the similarities of STS and F5 systems in the processing of actions. F5

cells may respond, in a manner analogous to the STS cells, to the sight of the agent reaching to grasp an object. The same F5 cells are active when the experimenter places an object behind a screen and then reaches as if to grasp it (even though the object is hidden from view during part of the action; Umilta et al., 2002). The sight of equivalent reaching when there is no reason to believe an object is hidden from sight fails to activate the F5 cells. Thus F5 and STS cells code the sight of actions on the basis of what is currently visible and on the basis of the recent perceptual history (Jellema & Perrett, 2002, 2003a; Jellema et al., 2002).

Thus, the visual properties of mirror neurons in F5 are strikingly similar to those described in STSa. Both F5 and STSa cells respond when the monkey observes the experimenter reaching and grasping an object, but will not respond to the sight of the experimenter's hand motion alone or to the sight of the object alone. In addition, the F5 cells respond to the corresponding sound of actions (Kohler et al., 2002). STS cells too appear to be sensitive to the sound of actions (Barraclough et al., 2003). In addition, the F5 cells but apparently not STS cells respond to the execution of the corresponding motor act. Therefore the F5 mirror system may provide a supramodal conceptual representation of actions and their consequences in the world. To some extent the polymodal sensory representation of actions may be constructed in the posterior and ventral regions of the brain (i.e., temporal lobe). The integration of the sensory representation with the motor representation of actions reflects an additional processing in parietal and frontal systems.

While STS and F5 cells have similar visual properties they may subserve distinct functions; the frontal system perhaps serves to control the behavior of the self particularly in dealing with objects (Rizzolatti et al., 1996a, 1996b), whereas the STS system is specialized for the detection and recognition of the behavior of others (Hietanen & Perrett, 1996; Mistlin & Perrett, 1990; Perrett et al., 1989).

The mirror neuron system might also complement the STSa description of the perceived action by adding information about the motor requirements of the perceived action, which could not easily be obtained from purely visual features. Crucially the properties of the frontal mirror system suggest that we may understand actions performed by others because we can match the actions we sense through vision and audition to our ability to produce the same actions ourselves. At a more speculative level, it has been proposed that the mirror neurons are involved in the ability to "read" others' minds. The cells may allow an observer to "experience" and understand an action performed by another through "simulating" that action (Gallese & Goldman, 1998).

Summary of STS Action Coding

Cells in the STSa of the macaque monkey code not only the sight of others' bodily actions but also combinations of other visual cues. These other visual cues derive either

from the body, for example, head/eye gaze direction and articulated body posture, or from the environment, for example, objects acted upon, and the spatial location of the actor. There are two ways in which the combined sensitivity arises: through spatial conjunctions with other cues that are simultaneously present, and through temporal conjunctions with other cues that are present consecutively.

Combining the sensitivity for actions with sensitivity for other bodily or environmental cues puts the STSa in a position to form representations of the causality or goal-directedness of others' actions. For instance, conjunctions of perceived actions with the perceived eye gaze direction of the agent performing the action may contribute to detecting the accidental or intentional outcome of the perceived action.

The relative spatial locations of the agent and the objects the agent interacts with (including the observer) may give insight to the agent's intention or goal. Populations of STSa cells are well equipped to keep track of such spatial relationships between agents and objects, which may further support comprehending the actions of others.

The object to which an action is directed is particularly important when it comes to interpreting goals. Some STSa cells are sensitive to the spatiotemporal interaction between the agent performing the action and the target-object of the action. Such selectivity ensures that the cells are optimally responsive in situations where the agent's motion is causally related to the object's motion.

The sensitivity to sequences might also contribute to representations of goals. The formation of associations between an articulated action and the static articulated end-posture of that action might well underlie the ability of the brain to imply impending or prior action from currently visible static postures. The performance of dexterous manual tasks can easily be specified as a series of static pictures, each demonstrating particular subgoals or stages in the action sequence. Based on an understanding of momentary postures during an action sequence, individuals can infer the dynamics of how an action was performed.

Other STS cells are tuned to the perceptual history within action sequences. Witnessing the history prior to viewing a static body allows one to predict the likelihood and nature of the body's future movement with more certainty than from a still image of a person performing a motor act. Natural actions are not isolated postures, but are continuous and complex sequences of postures with linking movements. The STS cells could play a role in recognizing complex action sequences and predicting the most likely next stage or consequence of actions.

Sequences of events are also crucial for those STS cells that code for agents hidden behind a screen. The observer needs to witness the agent disappearing behind the screen in order for the cell to produce a response to the hidden agent. The cellular responses to temporarily hidden agents combined with their sensitivity to direction of the agent's motion and spatial position allows again for prediction of when and where the agent may reappear.

Determining the goal or intention of an action not only involves the sensory systems but also emotional (Adolphs, 1999) and motor (Gallese & Goldman, 1998) systems. The ability to determine others' intentions is thus likely to be generated in a widely distributed network, involving many brain areas. The STSa could play a role in this network by providing descriptions of others' actions in terms of goals, intentions, or causes. These descriptions are, however, still mechanistic in nature. We have no evidence that descriptions include the attribution of mental states, such as motivational drives and beliefs (Baron-Cohen, 1994; Saxe et al., 2003), to the agent performing the action.

The advantages of being able to determine others' intentions from their actions are clear. The observer may anticipate the nature of the future actions of the other individual and adjust responses accordingly. An obvious question is of course whether there is behavioral evidence that nonhuman primates indeed discriminate between intentional and nonintentional actions (Byrne & Whiten, 1988). Macaque monkeys spontaneously follow the direction of attention of other monkeys (Emery et al., 1997; Lorincz, Baker, & Perrett, 1999). Similar gaze-following abilities have been described in different species of primates (see, for example, Tomasello, Call, & Hare, 1998). Attention-following does not, however, guarantee understanding of the consequences of attention, for example, that seeing leads to knowing. Studies by Call and Tomasselo (1998), and Hare and colleagues (2000, 2001) suggest that chimpanzees may understand what others can see, that seeing leads to knowing and the distinction between intentional and accidental actions. For monkeys, however, evidence of comprehension of beliefs or intentions is lacking.

In the following section we consider behavioral experiments designed to probe the monkey's understanding of an agent's actions and the possibility of anticipating impending actions on the basis of prior history and the agent's knowledge state.

Behavioral Studies of Actions, Intentions, and Beliefs

Although electrophysiology data tend to suggest that monkeys can anticipate others' actions and build some expectation of "what should come next" in a sequence of movements, behavioral data that support the existence of an even partial theory of mind in monkeys are scarce and have given negative results. For example, Hare et al. (2003) found that capuchin monkeys failed to understand what others can see in the same test in which chimpanzees had succeeded.

To get an insight into this issue, we examined the eye-gaze behavior of one macaque monkey as an index of the monkey's ability to predict the behavior of an experimenter in a feeding situation. We studied the gaze behavior as the monkey became accustomed to a particular feeding scenario. We reasoned that the monkey might learn to

predict the experimenter's actions and that the monkey's gaze could in principle show evidence of anticipating the intentions of the experimenter.

We simulated the "Sally-Anne" false belief test used with children. In this task, Sally places an object in one position, and leaves the room for a short time. During Sally's absence, Anne moves the object and hides it in a new location. Individuals with autism and typical children under four years incorrectly predict that Sally will search at the new shifted location on her return. Such individuals are said to lack a Theory of Mind because they do not understand that Sally's actions will be guided by her false belief that the object is where she left it.

Comprehension of a situation (physical or interpersonal) may outstrip actual overt behavioral performance. The understanding of the actions of others may develop faster than the ability to act on the basis of that understanding. Performance failure can occur because the operant behavior that is required by the experimental task is one that is guided by a pre-potent response. For example, chimpanzees are unable to withhold a reaching response to the more numerous (or larger) of two arrays of food rewards in a task where the experimental rule was to indicate the smallest array in order to get the biggest (Boysen & Berntson, 1995; Boysen, Berntson, Hannan, & Cacioppo, 1996; Boysen, Berntson, & Mukobi, 2001). If reaching is required not directly to one of two differently sized food rewards, but to one of two symbols (Arabic numerals) that represent the different quantities, then the chimpanzees were able to learn the rule to point to the symbol signifying the smaller quantity in order to receive the largest reward.

At a speculative level failure may reflect inadequate development of frontal cortex and an inadequate ability to give up short-term hedonistic impulses in exchange for the largest benefit that could come with deferred gratification. "Grab what you can now and don't care about what tomorrow might bring." Indeed, much of Old World monkeys' behavior might be guided by such short-term rules and not be subject to more strategic planning (Henzi & Barrett, 1999; Barrett & Henzi, 2000, 2002).

Similarly, several studies suggest that children under four years of age who fail traditional false-belief tasks may show a better understanding of false belief if they are tested with simplified procedures with less executive demands. For example, Carlson et al. (1998) found that children are better able to mislead an opponent if instead of pointing to an empty location they are allowed to mark it with a sticker. More pertinent for our study, Clements and Perner (1994) reported that children who fail a standard Sally-Anne false belief test with their verbal answers looked at the correct location.

Performance failures in tasks may therefore depend on the behavioral index chosen as an indication of comprehension. It may be that a reaching response, particularly for a reward, is too difficult to inhibit. By contrast, eye movements usually do not gain an individual reward. Where an individual looks may therefore be a better guide

to potential comprehension. Dissociations between knowledge expressed in gaze behavior and knowledge expressed in action have been found in other areas of cognitive development, such as object understanding (see Hauser, 2003, for a review).

Whatever the explanation, it is apparent that eye movements may provide the first indication of a developing Theory of Mind. The duration of looking is often used as an index of the ability of an individual to perceive and understand the difference between two stimuli or situations. It is assumed that the tested individual will look longer at the more unusual of the two stimuli, since habituation leads to decreased looking time and novel situations lead to longer durations of looking. Looking behavior can be compared in two scenarios in the Sally-Anne task. In one scenario Sally returns, goes, and searches at the site where she left the item (i.e., where her behavior would be guided by the nature of her false belief). In the second scenario Sally returns and goes and searches at the new site, the site where Anne has hidden the object. Here Sally's behavior is exceptional since she has a knowledge state that is inconsistent with her actual behavior.

Santos and Hauser (Hauser, 1999; Santos & Hauser, 1999 and personal communication) have studied the looking behavior of tamarin monkeys. Tamarins looked longer in the second scenario than in the first Sally-Anne scenario (Hauser, 1999; Santos & Hauser, 1999 personal communication). This would indicate that they perceive the difference between the two situations, but does not necessarily indicate that the monkeys have a Theory of Mind. To conclude that the monkeys can comprehend the actions of Sally based on Sally's false belief requires further tests. For example, if Sally remains present in the room while Anne shifts the location of the object, then Sally would have a true belief about the object's new location. In this case, when Sally attempts to retrieve the object, Sally should search at the new location, not at the location she originally placed the object.

Using an eye-tracking system, we studied the looking behaviour of one monkey in a Sally-Anne situation. The testing situation involved two boxes to the left and right of a central screen (1.02 m from the monkey). Boxes were oriented so that the monkey could always see the contents. By contrast Sally could not see the contents during her approach. On a "Start" trial, an experimenter (Sally, who normally feeds and interacts with the monkey) would enter the room from one side. There was then a period (average 3, range 2–5 seconds) during which Sally walked toward the monkey while hidden from sight behind the central screen. Sally then emerged on one side of the screen carrying food and a feeding device such as a spoon, which she would use to feed the monkey. Sally would then exit the scene on the same side as she entered, leaving the spoon in the box on that side. On a following "Anticipation" trial, Sally would always reenter the room with more food. Sally's approach was again obscured from sight during her return for 3 seconds behind the screen, after which she emerged to retrieve the spoon from where she had placed it and then fed the monkey.

During Start trials the spoon could be placed in the left or right boxes, but on ensuing anticipation trials Sally always returned to the side where she had placed the spoon. We measured the monkey's gaze during the period in which Sally was hidden from view behind the screen. We found that over sessions the monkey's eyes became governed by looking strategies. The first strategy was to look on the side of the screen Sally had been "*last seen*" before she became occluded behind the screen. This "*last seen*" strategy was evident on Start trials where there could be no prediction of the side on which the monkey was next going to be fed. Indeed the "*last seen*" strategy was superstitious, in that the side of Sally's actual reemergence following occlusion was random on such Start trials. Thus the monkey's anticipatory looking was not guided by any auditory cues to the approach or reappearance of Sally.

On Anticipation trials two strategies governed the anticipatory looking behavior during occlusion from sight: the "*last seen*" strategy and a new strategy which predicted Sally's actual behavior. The monkey would look more on the side Sally had previously left the feeding tool. This was the side from which "*Sally last fed*" the monkey and from which Sally would soon reemerge to collect the feeding tool and use it to feed the monkey.

This baseline behavior was established over 10 daily sessions (including 62 start and 220 anticipatory trials). In essence it showed that the monkey developed expectations about Sally's behavior in the feeding scenario. With this baseline established we progressed to "Critical trials." Here a second experimenter (Anne) entered after Anticipation trials and approached occluded from view behind the screen and then reemerged randomly on the left or right side and searched in the box for the spoon. If Anne did not find the spoon, she moved to the other box and searched there. Once Anne had found the spoon, she moved the spoon from the box from where Sally had placed it to the box on the opposite side. During all of this time Sally was away from the scene and therefore ignorant of the relocation of the spoon. On the next trial Sally returned and after the 3 seconds hidden approach behind the screen reemerged on the side where she had placed the spoon (where the spoon is according to her false belief). She searched there but did not find the spoon. She then searched the box on the opposite side, found the spoon and without pause returned to the original side to feed the monkey. In these critical trials the monkey's looking strategy (during Sally's temporary occlusion behind the screen) was significantly biased to the side where Sally should return guided by Sally's false belief (i.e., the side that did not contain the spoon).

We conducted various control trials to investigate whether the monkey's looking behavior was guided by Sally's knowledge state or some other rule. In "Critical Control" trials, Anne moved the spoon as before and exited. Anne then returned with food and after a 2–3 second pause hidden behind the screen, Anne emerged to search where she herself had left the spoon, which she picked up and then used to feed the

monkey. On these trials the monkey looked differentially at the scene during the time Anne was hidden. The monkey's eyes, however, did not anticipate where Anne would emerge (i.e., the monkey's fixation pattern was not guided by Anne's knowledge state). Indeed the monkey looked more to the side from which Anne had last been seen (*Anne last seen*) which could either be the side Sally last fed the monkey or the opposite side. On trials with Anne about to reappear, the monkey's gaze was governed by the prior visible behavior of Anne and not the prior behaviour of Sally. So the monkey's gaze and anticipation depended upon the identity of the experimenter.

The monkey's gaze was, however, *not* governed by Anne's knowledge (i.e., Anne's intention to collect the spoon and her "true belief" of its location). On more careful consideration (of all trial types, see below), Anne used the spoon to feed the monkey on only 17 percent of the times she was seen by the monkey to pick up the spoon. One can argue that the monkey had little reason to anticipate feeding by Anne, so the monkey's gaze behavior was perhaps appropriately unconnected with Anne's feeding actions.

We used two further control trials. In "Visible Swap" trials Sally witnessed Anne moving the spoon (here Anne has a true belief of the location of the spoon). In "touch trials" Anne touched but did not move the spoon. Combining the data from all unusual trial types (i.e., trial sequences involving Anne: Critical, Critical Control, Visible Swap, and Touch trials), we found that the monkey's pattern of looking most reflected the side on which Sally had last fed her rather than Sally's or Anne's knowledge state as to where the spoon was and hence where they would search (figure 10.2).

Thus the monkey's looking behavior revealed an understanding of the actions of others, but it was not governed by their knowledge state.

Nonetheless the monkey's looking behavior was subtly affected by the actions of Anne and Sally. We compared Critical trials (in which Sally is ignorant of Anne's behavior) with visible swap trials where Sally sees Anne's naughty tool-relocation behavior. This comparison revealed that the monkey's looking changed between the two trial types. In the Critical trials looking was governed solely by a *last fed* strategy that correctly anticipated Sally's reemergence, whereas in the Visible Swap trials looking was governed by both *last fed* and *last seen* strategies. Recall that the *last seen* strategy had been used on trials where Sally's impending behavior was unpredictable. This is tantalizing evidence that the monkey was less able to predict Sally's next step behavior on the basis of what had been witnessed (Anne relocating the tool and Sally seeing Anne doing this).

These studies were exploratory and the conclusions are tentative. We might be tempted to conclude that there was no evidence for Theory of Mind in monkeys. This may be premature because we do not yet know what humans with or without Theory of Mind would do in exactly the same control scenarios that we explored. Indeed, the studies of looking behavior by young children (Clements & Perner, 1994) and of

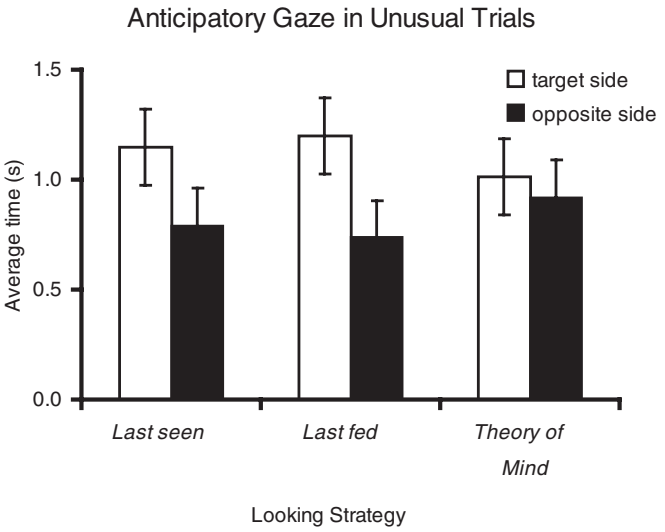


Figure 10.2
Gaze of one monkey during unusual trials in the Sally-Anne false belief scenario. Mean duration (sec) of anticipatory gaze to the target and the opposite side of a central occluding screen (where the target side was defined by the looking strategy). Gaze was measured in the 3–5 sec period during which the experimenter (Sally, or Anne) was hidden from sight prior to reappearing to search for a spoon to be used in feeding the monkey. The monkey's gaze was goverened overall by the strategy of looking to the side on which she had been *last fed* by Sally ($F_{1,40} = 5.0$, $p = 0.03$). There was a trend for looking to be guided by a *last seen* strategy (i.e., the side on which Sally had been seen prior to disappearance behind the screen, $F_{1,40} = 3.2$, $p = 0.08$). The *Theory of Mind* strategy, where looking would be guided by true or false beliefs of Sally (or Anne), did not predict looking ($F_{1,40} = 0.2$, $p = 0.6$).

tamarin monkeys (Hauser, 1999) concluded in favor of an understanding of belief on the basis of the same type of evidence we obtained in the Critical trials. Further studies are thus needed in human and nonhuman primates to investigate the strategies they use to guide their behavior in response to actions, intentions, and beliefs of others.

References

Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences*, *3*, 469–479.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*, *4*, 267–278.

Anderson, J. R., Montant, M., & Schmitt, D. (1996). Rhesus monkeys fail to use gaze direction as an experimenter-given cue in an object-choice task. *Behavioral Processes*, *37*, 47–55.

Ashbridge, E., Perrett, D. I., Oram, M. W., & Jellema, T. (2000). Effect of image orientation and size on object recognition: Responses of single units in the macaque monkey temporal cortex. *Cognitive Neuropsychology*, *17*, 13–34.

Baker, C. I., Keysers, C., Jellema, T., & Perrett, D. I. (2000). Coding of spatial position in the superior temporal sulcus of the macaque. *Current Psychology Letters: Behaviour Brain and Cognition*, *1*, 71–87.

Baker, C. I., Keysers, C., Jellema, T., Wicker, B., & Perrett, D. I. (2001). Neuronal representation of disappearing and hidden objects in temporal cortex of the macaque, *Experimental Brain Research*, *140*, 375–381.

Baron-Cohen, S., Ring, H., Moriarty, J., Schmitz, B., Costa, D., & Ell, P. (1994). Recognition of mental state terms. Clinical findings in children with autism and a functional neuroimaging study of normal adults. *British Journal of Psychiatry*, *165*(5), 640–649.

Barraclough, N. E., Xiao, D., Oram, M. W., & Perrett, D. I. (2003, 7–13 November 2003). Primate superior temporal sulcus (STS) neurons integrate visual and auditory information for biological motions. Paper presented at the 33th Annual Meeting of the Society For Neuroscience, New Orleans, LA.

Barrett, L., & Henzi, S. P. (2002). Constraints on relationship formation among female primates. *Behaviour*, *139*, 263–289.

Barrett, L., Henzi, S. P., Weingrill, T., Lycett, J. E., & Hill, R. A. (2000). Female baboons do not raise the stakes but they give as good as they get. *Animal Behaviour*, *59*, 763–770.

Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2003). Parallel visual motion processing streams for manipulable objects and human movements. *Neuron*, *34*, 149–159.

Boysen, S. T., & Berntson, G. G. (1995). Responses to quantity: Perceptual versus cognitive mechanisms in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology–Animal Behavior Processes*, *21*, 82–86.

Boysen, S. T., Berntson, G. G., Hannan, M. B., & Cacioppo, J. T. (1996). Quantity-based interference and symbolic representations in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology–Animal Behavior Processes*, *22*, 76–86.

Boysen, S. T., Berntson, G. G., & Mukobi, K. L. (2001). Size matters: Impact of item size and quantity on array choice by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, *115*, 106–110.

Bruce, C., Desimone, R., Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology*, *46*, 369–384.

Byrne, R. W., & Whiten, A. (1988). *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans*. Oxford: Clarendon Press.

Call, J., & Tomasselo, M. (1998). Distinguishing intentional from accidental actions in orangutans (*Pongo pygmaeus*), chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*). *Journal of Comparative Psychology*, *112*, 192–206.

Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The role of inhibitory processes in young children's difficulties with deception and false belief. *Child Development*, *69*(3), 672–691.

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, *9*, 377–395.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience*, *4*, 2051–2062.

Di Pellegrino, G., Fadiga, L., Fogassi, V., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, *91*, 176–180.

Dobbins, A. C., Jeo, R. M., Fiser, J., & Allman, J. M. (1998). Distance modulation of neural activity in the visual cortex. *Science*, *281*, 552–555.

Emery, N. J., Lorincz, E. N., Perrett, D. I., Oram, M. W., & Baker, C. I. (1997). Gaze following and joint attention in rhesus monkeys (*Macaca mulatta*). *Journal of Comparative Psychology*, *111*, 286–293.

Földiák, P. Xiao, D., Keysers, C., Edwards, R., & Perrett, D. I. (2004). Rapid serial visual presentation for the determination of neural selectivity in area STSa. *Progress in Brain Research*, *144*, 107–116.

Gallese, V., Fadiga, L, Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*, 593–609.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (2002). Action representation and the inferior parietal lobule. In W. Prinz & B. Hommel (Eds.), *Attention and Performance XIX* (pp. 334–355). Oxford: Oxford University Press.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *2*, 493–501.

Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature*, *349*(6305), 154–156.

Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology*, *35*, 96–111.

Hare, B., Addessi, E., Call, J., Tomasello, M., & Visalberghi, E. (2003). Do capuchin monkeys (*Cebus apella*) know what conspecifics do and do not see? *Animal Behaviour*, *65*, 131–142.

Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, *61*, 139–151.

Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, *59*, 771–785.

Hauser, M. D. (1999). Primate representations and expectations: Mental tools for navigating in a social world. In P. Zelazo, J. Astington, & D. Olson (Eds.), *Developing Theories of Intention: Social Understanding and Self-Control* (pp. 169–194). Hillsdale, NJ: Erlbaum.

Hauser, M. D. (2003). Knowing about knowing: Dissociations between perception and action systems over evolution and during development. *Annals of the New York Academy of Sciences*, *1001*, 79–103.

Henzi, S. P., & Barrett, L. (1999). The value of grooming to female primates. *Primates*, *40*, 47–59.

Hietanen, J. K., & Perrett, D. I. (1996). Motion sensitive cells in the macaque superior temporal polysensory area: Response discrimination between self- and externally generated pattern motion. *Behavioural Brain Research*, *76*, 155–167.

Iacoboni, M., Koski, L. M., Brass, M., Bekkering, H., Woods, R. P., Dubeau, M. C., Mazziotta, J. C., & Rizzolatti, G. (2001). Reafferent copies of imitated actions in the right superior temporal cortex. *Proceedings of the National Academy of Sciences 98*, 13995–13999.

Jellema, T., Baker, C. I., Oram, M. W., & Perrett, D. I. (2002). Cell populations in the banks of the superior temporal sulcus of the macaque and imitation. In W. Prinz & A. Meltzoff (Eds.), *The Imitative Mind: Development, Evolution, and Brain Bases* (pp. 267–290). Cambridge: Cambridge University Press.

Jellema, T., Baker, C. I., Wicker, B., & Perrett, D. I. (2000). Neural representation for the perception of the intentionality of actions. *Brain and Cognition*, *44*, 280–302.

Jellema, T., & Perrett, D. I. (2002). Coding of visible and hidden objects. In W. Prinz & B. Hommel (Eds.), *Common Mechanisms in Perception and Action: Attention and Performance XIX* (pp. 356–380). Oxford: O.U.P.

Jellema, T., & Perrett, D. I. (2003a). Perceptual history influences neural responses to face and body postures. *Journal of Cognitive Neuroscience*, *15*, 961–971.

Jellema, T., & Perrett, D. I. (2003b). Cells in monkey STS responsive to articulated body motions and consequent static posture: A case of implied motion? *Neuropsychologia*, *41*, 1728–1737.

Jellema, T., & Perrett, D. I. (2004). Single cell integration of animate motion, form, and location in the superior temporal sulcus of the macaque monkey. *Cerebral Cortex*, *14*, 781–790.

Kohler, E., Keysers, C., Umilta, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science*, *297*, 846–848.

Logothetis, N. K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Current Biology*, *5*, 552–563.

Lorincz, E. N., Baker, C. I., & Perrett, D. I. (1999). Visual cues for attention following in rhesus monkeys. *Current Psychology of Cognition*, *18*, 973–1001.

Mistlin, A. J., & Perrett, D. I. (1990). Visual and somatosensory processing in the macaque temporal cortex: The role of "expectation." *Experimental Brain Research*, *82*, 437–450.

Milner, A. D., & Goodale, M. A. (1995). *The Visual Brain in Action*. Oxford: Oxford University Press.

Murata, A., Gallese, V., Kaseda, M., & Sakata, H. (1996). Parietal neurons related to memory-guided hand manipulation. *Journal of Neurophysiology*, *75*, 2180–2186.

Nishitani, N., & Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences*, *97*, 913–918.

Nishitani, N., & Hari, R. (2001). Sign language and mirror neuron system. *Neuroimage*, *13*, S452.

Oram, M. W., & Perrett, D. I. (1994). Responses of anterior superior temporal polysensory (STPa) neurones to "biological motion" stimuli. *Journal of Cognitive Neuroscience*, *6*, 99–116.

Oram, M. W., & Perrett, D. I. (1996). Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *Journal of Neurophysiology*, *76*, 109–129.

Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., Chitty, A. J., Hietanen, J. K., & Ortega, J. E. (1989). Frameworks of analysis for the neural representation of animate objects and actions. *Journal of Experimental Biology*, *146*, 87–113.

Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London: Series B*, *335*, 23–30.

Perrett, D. I., Mistlin, A. J., Harries, M. H., & Chitty, A. J. (1990). Understanding the visual appearance and consequences of hand actions. In M. A. Goodale (Ed.), *Vision and Action: The Control of Grasping* (pp. 163–180). Norwood, NJ: Ablex.

Perrett, D. I., & Oram, M. W. (1993). Neurophysiology of shape processing. *Image and Vision Computing*, *11*, 317–333.

Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, *47*, 329–342.

Perrett, D. I., Smith, P. A. J., Mistlin, A. J., Chitty, A. J., Head, A. S., Potter, D. D., Broennimann, R., Milner, A. D., & Jeeves, M. A. (1985). Visual analysis of body movements by neurons in the temporal cortex of the macaque monkey: A preliminary report. *Behavioral Brain Research*, *16*, 153–170.

Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1984). Neurones responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity and relation to perception. *Human Neurobiology*, *3*, 197–208.

Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London: Series B*, *223*, 293–317.

Puce, A., & Perrett, D. I. (2003). Electrophysiology and brain imaging of biological motion. *Philosophical Transactions of the Royal Society of London*, *B*, *358*, 435–445.

Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., & Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey. II: Area F5 and the control of distal movements. *Experimental Brain Research*, *71*, 491–507.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu. E., Perani, D., & Fazio, F. (1996). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Experimental Brain Research*, *111*, 246–252.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*, 131–141.

Rizzolatti, G., & Gentilucci, M. (1988). Motor and visual-motor functions of the premotor cortex. In P. Rakic & W. Singer (Eds.), *Neurobiology of Neocortex* (*vol. 42*, 269–284). John Wiley & Sons.

Rolls, E. T., & Baylis, G. C. (1986). Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. *Experimental Brain Research*, *65*, 38–48.

Santos, L. R., & Hauser, M. D. (1999). How monkeys see the eyes: Cotton-top tamarins' reaction to changes in visual attention and action. *Animal Cognition*, *2*, 131–139.

Saxe, R., Xiao, D.-K., Kovacs, G., Perrett, D. I., & Kanwisher, N. (2004). Distinct representations of bodies, actions and thoughts in posterior superior temporal sulcus. *Neuropsychologia*, in press.

Sereno, A. B., & Maunsell, J. H. R. (1998). Shape selectivity in primate lateral intraparietal cortex. *Nature*, *395*, 500–503.

Tanaka, Y. Z., Koyama, T., & Mikami, A. (1999). Neurons in the temporal cortex changed their preferred direction of motion dependent on shape. *NeuroReport*, *10*, 393–397.

Tomasello, M., Call, J., & Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. *Animal Behaviour*, *55*, 1063–1069.

Umilta, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing: A neurophysiological study. *Neuron*, *31*, 155–165.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M.A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of Visual Behavior* (pp. 549–586). Cambridge, MA: MIT Press.

Wachsmuth, E., Oram, M. W., & Perrett, D. I. (1994). Recognition of objects, & their component parts: Responses of single units in the temporal cortex of the macaque. *Cerebral Cortex*, *5*, 509–522.

Williams, J. H. G., Whiten, A., Suddendorf, T., & Perrett, D. I. (2001). Imitation, mirror neurons, & autism. *Neuroscience, & Behavioural Reviews*, *25*, 287–295.

III Space, Action, and Attention: The Multiple Functions of Parietofrontal Circuits

11 The Mirror Neuron System and Its Role in Imitation and Language

Giacomo Rizzolatti and Giovanni Buccino

The human species is characterized by two fundamental cognitive abilities that are poorly developed or lacking altogether in other primates: imitation learning and the faculty of language. These abilities are obviously not the only ones that differentiate humans from other primates, including apes, but they certainly are among the most important.

Until recently it was hard to indicate, even in very speculative terms, the neural mechanisms underlying imitation and the faculty of language. This changed with the discovery of mirror neurons. The functional properties of these neurons suggest that primates are endowed with a mechanism that directly maps an observed action onto its motor counterpart. This matching mechanism appears to be able to provide a solution to the basic problem of imitation, that is, how an action described in visual terms may be replicated by the motor system, using completely different physiological parameters. The same mechanism may also give an account, although obviously highly speculative, on the riddle of language evolution.

We first summarize the functional properties of mirror neurons, the basic neural elements that constitute the mirror neuron system. We then compare the properties of human mirror neuron system with that of monkeys, examining in particular those properties that may account for the faculty of imitation in humans. In the last part of this chapter, we examine the relations between mirror neuron system and language.

Functional Properties of F5 Mirror Neurons

Mirror neurons were originally discovered in the rostral part of the ventral premotor cortex of the macaque monkey (area F5). Like all neurons of this area, mirror neurons have motor properties. They code mostly distal hand actions such as grasping, holding, tearing, and manipulating. Their defining functional characteristic is that they become active not only when the monkey does a particular action (like grasping an object) but also when it observes another individual (monkey or human) making

a similar action. Mirror neurons do not respond to the sight of a hand mimicking an action or to meaningless intransitive movements. Similarly, they do not respond to the observation of an object alone, even when it is of interest to the monkey (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Fogassi, & Gallese, 1996).

The vast majority of F5 mirror neurons show a marked similarity between the action effective in triggering them when observed and the action effective in activating them when executed. This sensory-motor congruence is occasionally extremely strict. In these cases the effective motor action and the effective observed action coincide both in terms of goal (e.g., grasping) and in terms of how the goal is achieved (e.g., precision grip). For most mirror neurons, however, the congruence is broader and is confined to the goal of the action.

Early studies on mirror neurons concerned essentially the upper sector of F5 where hand actions are mostly represented. Recently, a study was carried out on the properties of neurons located in the lower part of F5, where neuron activity is mostly related to mouth actions (Ferrari, Gallese, Rizzolatti, & Fogassi, 2003).

The results showed that about 25 percent of "mouth" neurons have mirror properties. According to the visual stimuli effective in triggering the neurons, two classes of mouth mirror neurons were distinguished: ingestive and communicative mirror neurons. Ingestive mirror neurons (80 percent of the recorded mouth mirror neurons) respond to the observation of actions related to ingestive functions (e.g., grasping food with the mouth). Virtually all of them show a good correspondence between the effective observed and the effective executed action.

More intriguing are the properties of the communicative mirror neurons. The most effective observed action is for them a communicative gesture such as lip smacking. However, as the ingestive mirror neurons, they strongly discharge when the monkey actively performs an ingestive action (figure 11.1).

This discrepancy between the effective visual input (communicative) and the effective active action (ingestive) is rather intriguing. Yet, there is evidence suggesting that, in evolution, communicative gestures, or at least some of them, derived from ingestive actions (see below). From this perspective one may argue that the communicative mouth mirror neurons found in F5 reflect a process of corticalization of communicative functions, not yet freed from their original ingestive basis.

An issue recently addressed was whether mirror neurons are able to recognize actions from their sound. Kohler and colleagues (2002) recorded F5 mirror neuron activity while the monkey was observing a "noisy" action (e.g., ripping a piece of paper), or was presented with the sound of the action without seeing it. The results showed that about 15 percent of mirror neurons responsive to presentation of actions accompanied by sounds also responded to the presentation of the sound alone. Most of them discharged specifically to the sound typical of the observed action. These neurons were dubbed "audiovisual" mirror neurons. The properties of audiovisual neurons strongly

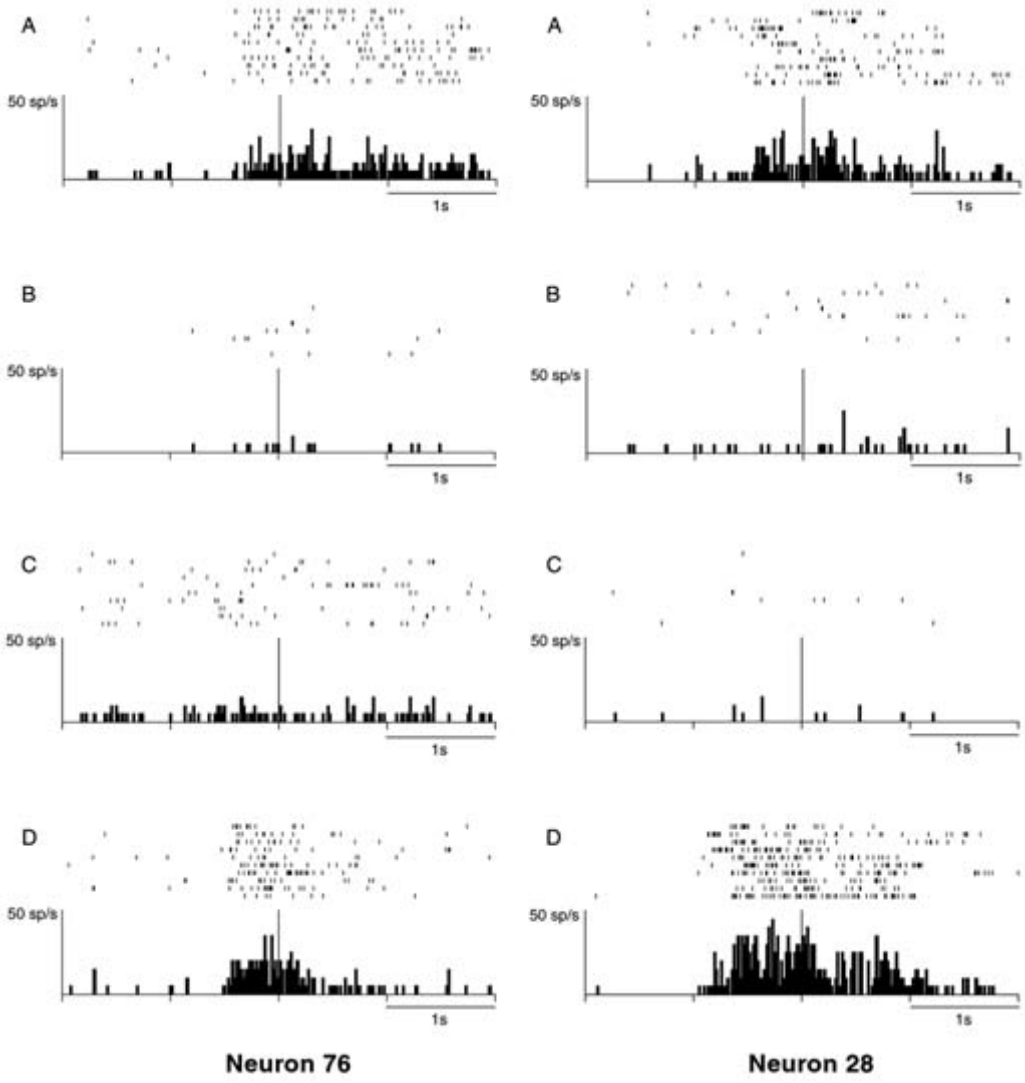


Figure 11.1

Examples of two communicative mirror neurons. *Left* (neuron 76): Activity during (*A*) observation of lip smacking, (*B*) observation of lips protruding, (*C*) observation of sucking a syringe. All gestures were made by an experimenter. (*D*) Discharge associated with monkey's active food ingestion accompanied by lips protruding. *Right* (neuron 28): Activity during (*A*) observation of lips protruding, (*B*) observation of holding food with the mouth, (*C*) food presentation. The gestures were made by an experimenter. (*D*) Discharge associate with monkey's active sucking juice from a syringe. Ordinates: spikes/s; abscissae: time, bin width 20 msec. (From Ferrari et al., 2003.)

suggest that the mirror neuron system is involved in action recognition, whatever the modality through which the action is presented.

Cortical Representation of Action Observation

Mirror neurons are not present only in area F5. A further cortical area where mirror neurons have been described is area 7 of Brodmann (1909) or area PF of von Economo (1929) (Fogassi, Gallese, Fadiga, & Rizzolatti, 1998; Gallese, Fogassi, Fadiga, & Rizzolatti, 2002). This area occupies the rostral part of the inferior parietal lobule. It receives input from the cortex buried in the superior temporal sulcus (STS) and sends an important output to ventral premotor cortex including area F5.

PF neurons are functionally heterogeneous. Most of them (about 90 percent) respond to sensory stimuli including complex visual stimuli such as biological motion. About 50 percent of PF neurons have, in addition, motor properties, discharging when the monkey performs specific movements or actions (Hyvarinen 1982; Fogassi et al., 1998, Gallese et al., 2002). Some of these neurons have clear mirror properties (Gallese et al., 2002).

Another cortical region in which neurons respond to the observation of actions done by others is the cortex of the superior temporal sulcus (STS) (Perrett et al., 1989; Perrett, Mistlin, Harries, & Chitty, 1990; Jellema, Baker, Wicker, & Perrett, 2000; see Jellema, Baker, Oram, & Perrett, 2002). Movements effective in eliciting neuron responses in this region are walking, turning the head, bending the torso, and moving the arms. A small set of STS neurons discharges also during the observation of goal directed hand-movements (Perrett et al., 1990).

If one compares functional properties of STS with F5 and PF mirror neurons, the major difference is that STS neurons do not discharge in association with monkeys' actions. STS neurons therefore do not appear to be endowed with motor properties.

In conclusion, the cortical mirror neuron circuit of the monkey is formed by two key regions: the rostral part of the inferior parietal lobule and the ventral premotor cortex. The functional significance of this system (mirror neuron system) will be discussed in the next paragraphs.

Function of the Mirror Neuron System in the Monkey: Action Understanding

There are two main hypotheses concerning the functional role of mirror neurons. The first is that mirror neuron activity underlies imitation (see Jeannerod, 1994), the second is that they are at the basis of action understanding (see Rizzolatti et al., 2001).

These two hypotheses are not mutually exclusive. However, because imitation is present, among primates, only in humans and (probably) in apes (see Galef 1988; Whiten & Ham, 1992; Byrne, 1995; Tomasello & Call, 1997; Visalberghi & Fragaszy,

2001), the evolutionary most ancient function of mirror neurons cannot be imitation. As it will be discussed later, imitation is a cognitive faculty that evolved later from the mirror neuron system following the acquisition of new matching properties by mirror neurons.

How do mirror neurons mediate understanding of actions done by others? The proposed mechanism is rather simple. Each time an individual sees an action done by another individual, neurons that represent that action are activated in the observer's premotor cortex. This automatically activated motor representation corresponds to that, which is spontaneously generated during active action and whose outcome is known to the acting individual. Thus, the mirror neuron system is able to transform visual information into knowledge (see Rizzolatti et al. 2001).

The Mirror Neuron System in Humans

A mirror neuron system also exists in humans. Evidence in this sense has been provided by a series of experiments carried out with various techniques such as electroencephalography (EEG), magnetoencephalography (MEG), transcranial magnetic stimulation (TMS), and brain imaging (see Rizzolatti et al., 2001).

These studies demonstrate that the human and monkey mirror neuron systems share the capacity to match observed actions onto their motor representations. However, they also showed some important differences.

First, in humans the observation of meaningful hand actions without an object (mimed actions) activates the mirror neuron system (Buccino et al., 2001; Grèzes, Armony, Rowe, & Passingham, 2003). The presence of an object appears to be necessary to activate the mirror neurons in the monkey (see above).

Second, TMS experiments show that there is a facilitation of the motor evoked potentials (MEPs) recorded from the muscles of the observer corresponding to those used by the actor performing the action. This facilitation is present also when an individual observes intransitive, meaningless hand/arm gestures (Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995; Iacoboni et al., 1999; Maeda, Kleiner-Fisman, & Pascual-Leone, 2002). Note that intransitive actions do not activate mirror neurons in the monkey.

Third, recent data show that, during action observation, there is not only a facilitation of MEPs recorded from the corresponding muscles, but also that the motor facilitation follows the time-course of the observed action (Gangitano, Mottaghy, & Pascual-Leone, 2001).

In conclusion, these properties indicate that the human mirror neuron system, unlike the monkey one, is able to describe both the goal of an action and the movements necessary to achieve it. The capacity to imitate not only the goal but also the way in which this is achieved is considered a necessary prerequisite for distinguishing true imitation from emulation and other pseudo-imitative behaviors.

Imitation

Imitation is often thought of as a cognitively undemanding, rather elementary form of behavior. Recent works across a variety of sciences show, however, that it is not true. There is clear evidence that imitation is a faculty particularly developed in humans, intrinsically linked to language and culture.

The involvement of the mirror neuron system in imitation was recently demonstrated by a series of brain imaging studies. Using the functional magnetic resonance imaging (fMRI) technique, Iacoboni et al. (1999) scanned normal human volunteers while they lifted a finger in response to: (1) the same action presented on a screen ("imitation"), (2) a symbolic cue, or (3) a spatial cue. The results showed that the activation was stronger during imitation than during the other motor conditions in the *pars opercularis* of the left inferior frontal gyrus (IFG), the right anterior parietal region, the right parietal operculum, and the right STS region (see also Iacoboni et al., 2001). Experiments by Koski et al. (2002) confirmed the importance of Broca's area, in particular when the action to be imitated had a specific goal. Grèzes et al. (2003) obtained similar results.

Nishitani and Hari (2000, 2002) performed two studies, using MEG, in which they investigated imitation of grasping actions and of facial movements, respectively. The first study confirmed the importance of the left IFG (Broca's area) in imitation. In the second study (Nishitani & Hari 2002), the authors asked volunteers to observe still pictures of verbal and nonverbal (grimaces) lip forms, to imitate them immediately after having seen them, or to make similar lip forms spontaneously. During lip forms observation, cortical activation progressed from the occipital cortex to the superior temporal region, the inferior parietal lobule, IFG (Broca's area), and finally to the primary motor cortex. The activation sequence during imitation of both verbal and nonverbal lip forms was the same as during observation.

In spite of some minor discrepancies, these data clearly show that the basic circuit underlying imitation coincides with that active during action observation. They also indicate that, in the posterior part of IFG, a direct mapping of the observed action and its motor representation takes place.

The importance of the *pars opercularis* of IFG in imitation was recently further demonstrated by Heiser et al. (2003), using repetitive TMS, a technique that transiently disrupts the functions of the stimulated area. The task used in the study was, essentially, the same as that of the fMRI study by Iacoboni et al. (1999). The results showed that following stimulation of both left and right Broca's area, there was significant impairment in imitation of finger movements. The effect was absent when finger movements were done in response to spatial cues.

Imitation Learning

In the experiments reviewed above individuals were asked to repeat on-line highly practiced actions made by another individual. A similar strategy was also used by Tanaka and Inui (2002), who asked volunteers to imitate on-line relatively complex hand or arm postures. In all these experiments the imitation consisted in matching the observed movements or actions to motor models already present in the parietal lobe and the premotor areas and to produce them. No motor learning was involved.

Buccino et al. (2004) recently addressed the issue of which cortical areas become active when individuals are required not simply to repeat an action or a posture present in their motor repertoire, but to produce, on the basis of action observation, a *novel motor pattern*. The basic task was imitation by naive participants of guitar chords played by an expert guitarist. By using an event-related fMRI paradigm, cortical activations were mapped during the following events: (1) observation of the chords made by an expert player, (2) pause (novel motor pattern formation and consolidation), (3) execution of the observed chords, and (4) rest. In addition to the imitation condition, there were three control conditions: observation of the guitar chords made by the player without any subsequent motor activity, observation of the chords followed by the execution of an actions not related to guitar chord execution (grasp-release of the guitar neck, rhythmical covering, or gentle scratching of the fretboard), free execution of guitar chords.

The results showed that during the event observation in the imitation condition there was activation of a cortical network formed by the inferior parietal lobule and the dorsal part of PMv plus the *pars opercularis* of IFG. This circuit was also active during the same event in the two control conditions in which participants merely observed the chords or observed them with the instruction to do subsequently an action not related to guitar chord execution. The strength of the activation was, however, much stronger during imitation then during the control conditions. During observation in the imitation condition there was, in addition, activation of the anterior mesial areas, superior parietal lobule, and a modest activation of the middle frontal gyrus.

The activations during the pause event in imitation condition involved the same basic circuit as in event observation, but with some important differences: increase of the superior parietal lobule activation, activation of PMd, and, most interestingly, a dramatic increase in extension and strength of the middle frontal cortex activation (area 46) and of the areas of the anterior mesial wall. Finally, during the execution event, the activation concerned, not surprisingly, the sensorimotor cortex contralateral to the acting hand.

These data show that the key centers for the formation of a novel motor pattern coincide with the key centers of the mirror neuron system. fMRI experiments do not give information on the mechanism underlying imitation; yet it is plausible (see the neurophysiological sections) that, during learning of new motor patterns by imitation, the observed actions are decomposed into elementary motor acts that activate, by mirror mechanism, the corresponding motor representations in PF and in PMv and in the *pars opercularis* of IFG. Once these motor representations are activated, they are recombined to fit the observed model. This recombination appears to occur inside the mirror neuron circuit with area 46 playing a fundamental orchestrating role.

Mirror Neuron System and Language

Some years ago Rizzolatti and Arbib (1998) proposed that the mirror matching mechanism represents the basic mechanism from which language evolved. This proposal was based on the consideration that mirror neurons create a direct link between the sender of a message and its receiver. Through them, therefore, observing and doing become manifestations of a single communicative (and, later, linguistic) faculty rather than two separate abilities.

Conceptually, the theory of Rizzolatti and Arbib belongs to theories that postulate that speech evolved, mostly, from gestural communication (see Armstrong, Stokoe, & Wilcox, 1995; Corballis, 2002). Its novelty consists in the fact that it indicates a neurophysiological mechanism that may create a common (parity requirement), nonarbitrary, link between the communicating individuals. This link can hardly be created by sounds. Sounds, by their nature, cannot generate the shared, nonarbitrary knowledge that can be achieved through the involvement of the motor system.

Humans mostly communicate by sounds. Thus, it seems almost natural to consider human speech as an evolutionary extension of the sound-based animal communication. In fact, human speech and animals' calls are different phenomena. First of all, the structures underlying speech and animals' calls in mammals are completely different. Animals' calls are mediated primarily by the cingulate cortex and by diencephalic and brain stem structures (see Jurgens, 2002). In contrast, human speech has its core substrate in the perisylvian areas, including area 44, a premotor area. Second, speech is not necessarily linked to an emotional behavior, whilst animals' calls are. Third, speech is mostly a person-to-person communication. In contrast animal calls are, typically, directed to "everybody," rather than to a specific individual. Fourth, but not least, speech is endowed with combinatorial properties that are absent in animal communication. It is recursive and virtually limitless with respect to its scope of expression.

The dichotomy between a communication system as that of animals' calls and the one that eventually led to speech is nicely described by Sir Richard Paget (1930). He

writes, "It may be imagined that, in the early stages of human development mankind roared and grunted and sung, on the one hand, to express his emotions, and gesticulated and grimaced on the other to explain his ideas. In some cases he may have used both methods together, as when the dog makes the threatening gestures of his teeth, and energizes or phonates this gesture by the addition of a laryngeal growl."

There is evidence that, in humans, some types of emotional communication are based on a specific type of mirror mechanism (Carr et al., 2003, Krolak-Salmon et al., 2003; Wicker et al., 2003). One may argue, therefore, that in terms of its basic mechanism, emotional system had the same potentiality to become a system conveying referential information as the system related to actions. The classical studies of the alarm calls of vervet monkeys, as well as other studies that extended the observation to other species and other communicative contexts like social relationship, food, intergroup aggression, show that evolution indeed tried this pathway. This attempt was doomed, however, to failure. As noted by Hauser et al. (2002), unlike the animal examples of referential signals, most of the words of human language are not linked to a specific function (e.g., a warning cry), but "can be linked to a multiplicity of concepts." In a nonemotional communication system, the same word (e.g., fire), may indicate that fire erupted ("escape" message), but it may also indicate that the fire is ready and we can prepare the meal ("approach" message) or an almost infinite series of other meanings. In contrast, in an emotion based communicative system, a call may have a referential meaning, but, essentially, has the function of starting a specific behavior. This link with a specific response renders the emotional system unsuitable for language evolution, in spite of its mirror mechanism.

Mirror Mechanism and Sign-Language Evolution

Monkey mirror neurons code object-directed actions. Within this limit, they solve two fundamental communication problems: parity and direct comprehension of the action. Parity requires that what counts for the sender of the message, counts also for the receiver. Direct comprehension means that there is no need of an agreement between individuals to understand each other. No arbitrary symbols are required. The comprehension is inherent to neural organization of the individual.

The monkey mirror neuron system is, however, a closed system linked to objects. A first problem for the mirror neuron theory of language evolution is to explain how this close object-related system became an open system, able to describe actions and objects without directly referring to them.

It is likely that the great leap from an object-related mirror neuron system to a truly communicative mirror neuron system is related to the development of imitation (see Arbib, 2002) and the related changes observed in the human mirror neuron system:

the capacity of mirror neurons to respond to pantomimes (Buccino et al., 2001; Grèzes et al., 2003) and to discharge in response to intransitive actions (Fadiga et al., 1995; Maeda, Kleiner-Fisman, & Pascual-Leone, 2002).

It is possible that these modifications of the mirror neuron system did not evolve originally in order to communicate, but resulted as a consequence of the necessity to learn, by imitation, actions made by others. Imitation implies not only the understanding of the purpose of the action to be imitated, but also the capacity to repeat the single movements that constitute an action in the right order (Tomasello & Call, 1997; Rizzolatti, 2005). The necessity to keep trace of precise movements sharpened the mirror neuron system and its capacity to convey information.

The idea that communicative actions derive from other evolutionary more ancient actions is not new. Van Hoof (1967), for example, proposed it in his work on the origin of monkey communicative gestures. According to him, many of the most common communicative gestures, such as lip-smacking or lips-protruded face, are ritualizations of ingestive actions that monkeys use for affiliative purposes. Similarly, MacNeilage (1998) suggested that the human vocal communication derived from the cyclic open-close mandibular alternation originally evolved for food ingestion. According to this view, monkey lip-smacking represents a communicative action derived from ingestive mouth movements. The existence of a neurophysiological link between ingestive and communicative actions is confirmed by the properties of F5 mouth mirror neurons (Ferrari et al., 2003; see above).

A similar notion is held for action development by Vygotsky (1934), who proposed that pointing derives from children attempts to grasp objects. When objects are located close to a child, the child grasps them, while, when they are located far from it, the child extends its arm and hand toward. An object-related action becomes an intransitive communicative action.

Taken together, these findings suggest that in humans (and most likely in human ancestors as well), pantomimes of actions and a variety of intransitive actions were incorporated into the mirror neuron system, thus acquiring its communicative properties.

We agree with Arbib (2002) that the gestural phase of communication, here just sketched, did not reach the sophisticated complexity of the modern sign language. It plausible that "protosigns" were soon accompanied by sounds and that speech development prevented the occurrence of a full-fledged sign language. The protosign language allowed, however, individuals to communicate in a much richer way than it was possible by using the emotional system. Protosigns allowed individuals to describe directions and action locations, to pantomime actions, and to give iconic descriptions of objects. Protosign language should have given a strong evolutionary advantage to individuals able to use it, providing in this way a strong stimulus for further evolution of communication.

Mirror Neurons and Speech Evolution

The protosign communication system has a great asset: its semantics is neither arbitrarily imposed nor derives from an improbable agreement among individuals. It is inherent to the gestures that are used to communicate. This is not so, or at least is not apparent, for speech. Indeed, one of the most difficult aspects of speech evolution is to provide a satisfactory answer to the problem of how words started to signify things.

Historically, the discussion on how this occurred is centered on the possible relations between the sound of a word and its meaning. On one side, there are those who postulate a remote, but "natural" origin of the words, on the other those who regard the faculty of speech "with almost superstitious veneration, and, emulating the etymologists of Socrates, are content to ascribe the first words to arbitrary choice of gods" (Critchley, 1939, p. 15). Even if "gods" are substituted with a more prosaic concept as general agreement among speakers, the hypothesis of a natural origin of language seems to be the only one intellectually satisfactory.

A major problem with the "natural" theory, however, is the difficulty of specifying what is the evolutionary link between the sound of a word and its meaning. Onomatopeia, that is the similarity between sound of the words and the noise produced by natural events or actions, done by humans or animals, is one of the suggested possibilities. Another possibility is represented by interjectional utterances emitted by individuals in certain conditions. The problem with both these hypotheses is that they are able to explain a very limited number of words. Thus, although they trace the origin of some words, they lack the generality necessary to explain the link between sound and word for most of them.

An interesting alternative hypothesis was advanced by Paget (1930). According to him, the original human communication was gestural. However, as the individual gesticulated with his hand, "his tongue, lips and jaw unconsciously followed suit in a ridiculous fashion, understudying the action of the hands." Later, the gesticulating individuals discovered that the expiration of the air through the oral cavities produced audible gestures, that is, voiced speech.

Paget gives many examples of parallelism between sound and meaning in a variety of languages. For example, as far as vowels are concerned, he suggests that "A" (as in large) refers to anything that is large, wide open, spacious; "I" (as in mini) to something that is small or pointed; "AW" connotes a cavity (e.g., yawn) and "OO" something tubular or elongated. Consonants also convey gestural symbolisms. "M" implies a continued closure; "R" implies a bending back; "DR" and "TR" denote running or walking, the direction of the tongue movement being inwards toward the speaker.

According to this theory, called "schematopoeia," the great majority of words appear to be pantomimic. They are built "much as the Chinese ideographs are, by addition of separately significant elements" (Paget, 1930). This type of organization explains why it is difficult to discover the original sound meaning in the words. Furthermore, almost every action or idea can be pantomimed in many different ways and every gesture can be construed in many different ways. This account for another aspect of speech that was considered evidence against a "natural" origin of language: the dissimilarity among languages. Yet, in spite of this, similar word can be found in unrelated languages such archaic Chinese and Sumerian and other words that are dissimilar can be deciphered using the basic semantic values given to vowels and consonants.

It is obvious that the schematopoeia theory is essentially a speculation. Yet, its basic notion that the hand/body gestures and the primitive speech gestures were intrinsically linked is very interesting. On one hand, it provides a possible clue on how intrinsically known messages (hand gestures) were transferred to an opaque gestural system, as the orolaryngeal system, without loosing their intrinsic (nonarbitrary) meaning. On the other hand, a clear neurophysiological prediction derives from it: hand/arm and speech gestures must be strictly linked and have, at least partially, a common neural substrate.

A series of recent studies demonstrates that this prediction is true. TMS experiments showed that the right-*hand* motor excitability increases during reading and spontaneous speech (Tokimura et al., 1996; Seyal et al., 1999; Meister et al., 2003). The effect is limited to the left hemisphere. No language-related effects are found in the motor area of the leg. Meister et al. (2003) stressed that the increase of hand motor cortex excitability cannot be due to word articulation, because while word articulation recruits motor cortex bilaterally, the observed activation was strictly limited to the left hemisphere. The facilitation appears, therefore, to result from a coactivation of the right hand motor area with the language network.

Gentilucci and coworkers (2001) reached similar conclusions using a completely different approach. In a series of behavioral experiments, they asked participants to grasp two objects of different size with their mouths and, simultaneously, to open their right hands. The results showed that the maximal finger aperture and time to maximal finger aperture increased when the mouth was directed to the large object.

Even more relevant to the view of a strict link between hand actions and orolaryngeal gestures is another experiment of the same study (Gentilucci et al., 2001). Participants were presented with two 3-D objects, one large, and the other small. On the visible face of the objects either two symbols or a series of dots randomly scattered on the same area occupied by the symbols were written. Participants were required to grasp the objects, but, in the condition in which the symbols appeared on

the object, they had to open their mouths. The kinematics of hand, arm, and mouth movements was recorded. The results showed that, although participants were instructed to keep the mouth aperture constant in all conditions, lip aperture and the peak velocity of lip aperture increased when the movement was directed to the large object. Control experiments showed that the effect was specific to movements of the mouth and of the contralateral hand movements. Simultaneous extension of the contralateral forearm was not affected by the main task.

In a further experiment the same authors adopted the same experimental procedure described above, but asked the participants to pronounce a syllable (e.g., GU, GA) instead of simply opening their mouths. The syllables were written on the object in the same location where the symbols appeared in the previous experiment. It was found that lip aperture was larger when the participants grasped a larger object. Furthermore, the maximal voice power recorded during syllable emission was also higher when grasping the larger object (Gentilucci, 2001).

It is clear from these experiments that both simple buccal movements and the orolaryngeal synergies necessary for syllable emission are linked to manual gestures. Most importantly, hand actions requiring large movements share neural organization with orolaryngeal movements coding large mouth movements. This is reminiscent of the claim that vowel A describes something large, while the vowel I (EE) describes something small.

Grasping movements influence syllable pronunciation not only when executed but also when observed. Participants were required to pronounce the syllables BA or GA while observing another individual grasping objects of different size. It was found that the kinematics of lip aperture and the amplitude spectrum of voice were influenced by the observation of grasping movements done by another individual. Specifically, both lip aperture and voice peak amplitude were greater when the observed action was directed to larger objects. Control experiments ruled out that the effect was due to the size of the object or to the velocity of the observed arm movement (Gentilucci, 2003).

Finally, evidence for a link between gesturing and speech system also comes from clinical studies. Hanlon, Brown, and Gerstman (1990) showed that, in aphasics, pointing with the right hand to a screen where objects are presented facilitates object naming. Similarly, Hadar and colleagues (1998) found that word retrieval is facilitated through gesturing in brain-damaged patients.

It is obvious that the reviewed experiments by no means prove the schematopoeia theory. Yet, they indicate (those of Gentilucci in particular) that the theory is not so bizarre as one may think and that the link between hand gestures and speech system is extremely strong also in the extant *Homo sapiens*.

Auditory Mirror Neuron System

Let us accept that the meaning of manual gestures, "naturally" understood through the mirror neuron mechanism, transferred, at a certain point of evolution, to orolaryngeal gestures and that this transfer marked the beginning of language based on sound. Is this assumption plausible?

The presence of audiovisual mirror neurons in old-world monkeys (Kohler et al., 2002; see also above) suggests that auditory access to action representation is a feature common to many primate species. It is likely, therefore, that a link between auditory stimuli and action representation was also present in the primates from whom *Homo sapiens* descended. Thus, before speech occurrence, the precondition for the transfer of gesture meaning from visual to auditory modality was already present.

The monkey audiovisual mirror neurons, however, code only object-related actions. Their function is similar, with the addition of auditory responses, to "classical" mirror neurons. These neurons, however, as discussed above, are not sufficient to create an efficient communicative system. To achieve such communication, meaningful sounds (and later in evolution words) should stem from the association of sounds with intransitive actions and pantomimes done with the mouth and with the arms.

We already discussed some possibilities on how this visuoauditory transfer may have occurred. Let us examine now its consequences for the cortical organization. An example may serve to this purpose.

When we eat, we move our mouth, tongue, and lips in a specific manner. The observation of this combined series of motor actions constitutes a gesture whose meaning is transparent: "eat." If, while making this action, we blow air through the orolaryngeal cavities, we produce a sound like "mnyam-mnyam," or "mnya-mnya," words whose meaning is almost universally recognized (Paget, 1930). Such a mechanism allows the transfer of the meaning of an action, "naturally" understood, to a sound.

The understanding of words related to orolaryngeal gestures (like "mnyam-mnyam"), should have initially occurred through activation of audiovisual mirror neurons related to ingestive behavior or to mouth movements accompanying hand gestures. The fundamental step toward speech acquisition was achieved, however, only when individuals, thanks to their improved imitation capacities, became free to generate the sounds of actions without actually performing those actions. In analogy to what occurred to the classical mirror neuron system, this evolutionary step should have been accompanied by the acquisition, by the motor neurons controlling oropharyngeal gestures, of the capacity to resonate to sound emitted by a similar motor gestures. Thus, a new type of mirror neuron controlling sound emission and responding to meaningful sound should have been generated ("echo mirror neurons"). The incredibly confused organization of human Broca's area, where phonology, semantics,

hand actions, ingestive actions, and syntax are all intermixed in a rather restricted neural space (Bookheimer, 2002) is probably a consequence of this evolutive trend.

Is there any evidence that humans possess an echo mirror neuron system, that is, a system that motorically "resonates" when the individual listen to verbal material? Recent evidence shows that this is the case.

Fadiga et al. (2002) recorded MEPs from the tongue muscles in normal volunteers instructed to listen carefully to acoustically presented verbal and nonverbal stimuli. The stimuli were words, regular pseudo-words, and bitonal sounds. In the middle of words and pseudo-words either a double "f" or a double "r" were embedded. "F" is a labiodental fricative consonant that, when pronounced, requires slight tongue movements, while "r" is linguopalatal fricative that, in contrast, requires, a marked tongue muscles involvement to be pronounced. During the stimulus presentation the participants' left motor cortex was stimulated with single pulse TMS.

The results showed that listening to words and pseudo-words containing the double "r" determines a significant increase of MEPs' amplitude recorded from tongue muscles with respect to listening to bitonal sounds, and words and pseudo-words containing the double "f." Furthermore, the facilitation due to the listening of the "r" consonant was stronger for word than for pseudo-words (figure 11.2).

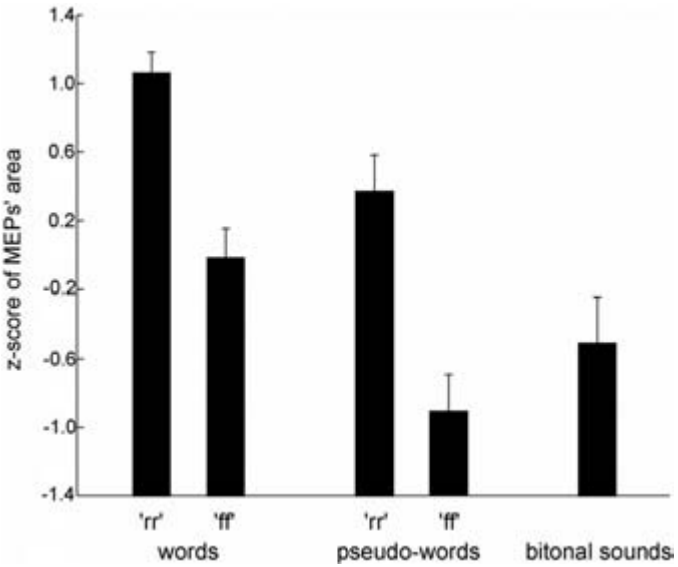


Figure 11.2

Normalized motor evoked potential (MEPs) areas recorded from tongue muscles during listening to words, pseudowords and bitonal sounds. Data from all subjects. "rr" refer to verbal stimuli containing lingual-palatal fricative consonants; "ff" refer to verbal material containing labiodental fricative consonants. (From Fadiga et al., 2002.)

Results congruent with those of Fadiga were obtained by Watkins et al. (2003). By using single pulse TMS technique, they recorded MEPs from a lip (*orbicularis oris*) and a hand muscle (first *interosseus*), respectively, in four conditions: listening to continuous prose, listening to nonverbal sounds, viewing speech-related lip movements, and viewing eye and brow movements. Compared to control conditions, listening to and viewing speech enhanced the MEPs' amplitude recorded from the *orbicularis oris* muscle. This increase was seen only in response to stimulation of the left hemisphere. No changes of MEPs in any condition were observed following stimulation of the right hemisphere. Finally, the size of MEPs elicited in the first *interosseus* muscle did not differ in any condition.

Taken together, these data undoubtedly show that an echo mirror neuron system exists in humans: when an individual listen to verbal stimuli there is an automatic activation of the speech-related motor centers. The precise functional meaning of this system and its role in word understanding is, however, more difficult to specify.

A first possibility is that the echo mirror neuron system evolved exclusively for production/imitation of verbal sounds and still is used only for this purpose. Another hypothesis, not in contrast with the first one, is that the echo mirror neuron system subserves, besides imitation, speech "perception" ("motor theory of speech perception," Liberman et al., 1967; Liberman & Mattingly, 1985; Liberman & Whalen, 2000). According to this theory, the ultimate constituents of speech are not sounds but articulatory gestures that evolved for the service of language. From this perspective, the echo mirror neuron system would represent the neural mechanism transforming verbal sounds into the motor representation of the corresponding articulatory gestures.

It is important to note that the understanding of word semantics is not part of Liberman's theory. On the contrary, the evolutionary scenario, sketched above, predicts also a semantic role for the echo mirror neurons. According to it, there are two roots to semantics: one, more ancient, based on the auditory (verbal material) activation of mirror neurons coding actions (see Pulvermueller, 2002), the other, evolutionary more recent, based on the activation of echo mirror neurons.

This second root implies a rather interesting concept: a *second-order action representation*. Classical mirror neurons do not require action execution in order to be triggered. The observed action is understood without its actual execution. The mirror neurons activation *represents* by itself the action (*first-order action representation*). We propose that a similar process takes place during the activation of echo mirror neurons.

The hypothesis is the following. The echo-mirror neurons become active in response to verbal material and their activation evokes the motor representation of the corresponding articulatory gestures. Thus, in the case of "mnyam-mnyam," this sound recruits the corresponding articulatory representation. There is, however, something

more. Because the echo mirror neurons (e.g., those activated by the "mnyam-mnyam" sound), given their evolution origin must be connected with the classical mirror neurons coding the corresponding actions (e.g., to eat), when they discharge the perceiving individual recognizes not only the correct sound of the word but also its meaning.

In more general terms, at least in a competent speaker, the activity of the echo mirror neurons is sufficient by itself to represent the content of mirror neurons linked to them, without the absolute necessity of mirror neuron system activation (and even less so of motor neurons controlling action execution). The echo-mirror neuron activation *represents*, therefore, albeit indirectly, an action (*second order action representation*). This second-order representation, located in a circumscribed neural space, has clear advantage over the first-order representation. It has the power to create, for example, new associations between words based on the probability of word occurrence rather than on the occurrence of actions and gives, in this way, much higher communicative possibilities to an individual possessing it than those achievable with the ancient semantic system based on the evocation of motor representations.

It is obvious that simple words, like the one discussed in the previous example, are rather rare, and a more complex system associating phonology to semantics should have evolved to link the sound of polysyllabic words (and the grammatical variations of all words) to meaning. We think, however, that the basic idea, here presented, that the echo mirror neuron system represents a link between phonology and semantics has a strong heuristic value, and, being based on neurophysiological findings, can undergo experimental testing, and possibly bridge linguistic data with neurophysiology.

References

Arbib, M. A. (2002). Beyond the mirror system: Imitation and evolution of language. In K. Dautenhan & C. Nehaniv (Eds.), *Imitation in Animals and Artifacts* (pp. 229–280). Cambridge, MA: MIT Press.

Armstrong, A. C., Stokoe, W. C., & Wilcox, S. E. (1995). *Gesture and the Nature of Language*. Cambridge: Cambridge University Press.

Bonin, G. von, & Bailey, P. (1947). *The Neocortex of Macaca mulatta*. Urbana: University of Illinois Press.

Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Reviews in Neuroscience*, *25*, 151–188.

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien Dargerstellt auf Grund des Zellenbaues*. Leipzig: Barth.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience*, *13*, 400–404.

Buccino, G., Vogt, S., Ritzl, A., Fink, G. R., Zilles, K., et al. (2004). Neural circuits underlying imitation of hand actions: An event related fMRI study. *Neuron*, in press.

Byrne, R. W. (1995). *The Thinking Ape: Evolutionary Origins of Intelligence*. Oxford: Oxford University Press.

Calder, A. J., Keane, J., Manes, F., Antoun, N., Young A. W. 2000. Impaired recognition and experience of disgust following brain injury. *Nature Neuroscience*, *3*, 1077–1088.

Carr, L., Iacoboni, M., Dubeau, M. C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences*, *100*, 5497–5502.

Corballis, M. C. (2002). *From Hand to Mouth: The Origins of Language*. Princeton, NJ: Princeton University Press.

Critchley, M. (1939). *The Language of Gesture*. London: Edward Arnold.

Economo, C. von. (1929). *The Cytoarchitectonics of the Human Cerebral Cortex*. London: Oxford University Press.

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, *15*, 399–402.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology*, *73*, 2608–2611.

Ferrari, P. F., Gallese, V., Rizzolatti, G., & Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience*, *17*, 1703–1714.

Fogassi, L., Gallese, V., Fadiga, L., & Rizzolatti, G. (1998). Neurons responding to the sight of goal directed hand/ arm actions in the parietal area PF (7b) of the macaque monkey. *Society for Neuroscience Abstracts 24*, 257.5.

Galef, B. G. (1988). Imitation in animals: History, definition and interpretation of data from psychological laboratory. In T. Zental & B. G. Galef (Eds.), *Comparative Social Learning* (pp. 3–28). Hillsdale, NJ: Erlbaum.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*, 593–609.

Gallese, V., Fogassi, L., Fadiga, L., & Rizzolatti, G. (2002). Action representation and the inferior parietal lobule. In W. Prinz & B. Hommel (Eds.), *Attention & Performance XIX: Common Mechanisms in Perception and Action* (pp. 335–355). Oxford: Oxford University Press.

Gangitano, M., Mottaghy, F. M., & Pascual-Leone, A. (2001). Phase specific modulation of cortical motor output during movement observation. *NeuroReport*, *12*, 1489–1492.

Gentilucci, M. (2003). Grasp observation influences speech production. *European Journal of Neuroscience*, *17*, 179–184.

Gentilucci, M., Benuzzi, F., Gangitano, M., & Grimaldi, S. (2001). Grasp with hand and mouth: A kinematic study on healthy subjects. *Journal of Neurophysiology*, *86*, 1685–1699.

Grèzes, J., Armony, J. L., Rowe, J., & Passingham, R. E. (2003). Activations related to "mirror" and "canonical" neurons in the human brain: An fMRI study. *Neuroimage*, *18*, 928–937.

Hadar, U., Wenkert-Olenik, D., Krauss, R., & Soroker, N. (1998). Gesture and the processing of speech: Neuropsychological evidence. *Brain and Language*, *62*, 107–126.

Hanlon, R. E., Brown, J. W., & Gerstman, L. J. (1990). Enhancement of naming in nonfluent aphasia through gesture. *Brain and Language*, *38*, 298–314.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, *22*(298), 1569–1579.

Heiser, M., Iacoboni, M., Maeda, F., Marcus, J., & Mazziotta, J. C. (2003). The essential role of Broca's area in imitation. *European Journal of Neuroscience*, *17*, 1123–1128.

Hyvarinen, J. (1982). Posterior parietal lobe of the primate brain. *Physiological Review*, *62*, 1060–1129.

Iacoboni, M., Koski, L. M., Brass, M., Bekkering, H., Woods, R. P., et al. (2001). Reafferent copies of imitated actions in the right superior temporal cortex. *Proceedings of the National Academy of Sciences*, *98*, 13995–13999.

lacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, *286*, 2526–2528.

Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, *17*, 187–245.

Jellema, T., Baker, C. I., Wicker, B., & Perrett, D. I. (2000). Neural representation for the perception of the intentionality of actions. *Brain and Cognition*, *442*, 280–302.

Jellema, T., Baker, C. I., Oram, M. W., & Perrett, D. I. (2002). Cell populations in the banks of the superior temporal sulcus of the macaque monkey and imitation. In A. N. Melzoff & W. Prinz (Eds.), *The Imitative Mind: Development, Evolution and Brain Bases* (pp. 267–290). Cambridge: Cambridge University Press.

Jurgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience Biobehavior Review*, *26*, 235–258.

Kohler, E., Keysers, C., Umilta, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science*, *297*, 846–848.

Koski, L., Iacoboni, M., Dubeau, M. C., Woods, R. P., & Mazziotta, J. C. (2003). Modulation of cortical activity during different imitative behaviors. *Journal of Neurophysiology*, *89*, 460–471.

Koski, L., Wohlschlager, A., Bekkering, H., Woods, R. P., & Dubeau, M. C. (2002). Modulation of motor and premotor activity during imitation of target-directed actions. *Cerebral Cortex*, *12*, 847–855.

Krolak-Salmon, P., Henaff, M. A., Isnard, J., Tallon-Baudry, C., Guenot, M., Vighetto, A., Bertrand, O., & Mauguiere, F. (2003). An attention modulated response to disgust in human ventral anterior insula. *Annals of Neurology*, *53*, 446–453.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*, 431–461.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*, 1–36.

Liberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Neuroscience*, *4*, 187–196.

MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences*, *21*, 499–511.

Maeda, F., Kleiner-Fisman, G., & Pascual-Leone, A. (2002). Motor facilitation while observing hand actions: Specificity of the effect and role of observer's orientation. *Journal of Neurophysiology*, *87*, 1329–1335.

Meister, I. G., Boroojerdi, B., Foltys, H., Sparing, R., Huber, W., & Topper, R. (2003). Motor cortex hand area and speech: Implications for the development of language. *Neuropsychologia*, *41*, 401–406.

Nishitani, N., & Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences*, *97*, 913–918.

Nishitani, N., & Hari, R. (2002). Viewing lip forms: Cortical dynamics. *Neuron*, *36*, 1211–1220.

Paget, R. (1930). *Human Speech*. London: Kegan Paul, Trench.

Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., et al. (1989). Frameworks of analysis for the neural representation of animate objects and actions. *Journal of Experimental Biology*, *146*, 87–113.

Perrett, D. I., Mistlin, A. J., Harries, M. H., & Chitty, A. J. (1990). Understanding the visual appearance and consequence of hand actions. In M. A. Goodale (Ed.), *Vision and Action: The Control of Grasping* (pp. 163–342). Norwood, NJ: Ablex.

Pulvermueller, F. (2002). *The Neuroscience of Language*. Cambridge: Cambridge University Press.

Rizzolatti, G. (2005). The mirror neuron system and imitation. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science,* vol. 1 (pp. 55–76). Cambridge, MA: MIT Press.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neuroscience*, *21*, 188–194.

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*, 131–141.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, *2*, 661–670.

Seyal, M., Mull, B., Bhullar, N., Ahmad, T., & Gage, B. (1999). Anticipation and execution of a simple reading task enhance corticospinal excitability. *Clinical Neurophysiology*, *110*, 424–429.

Tanaka, S., & Inui, T. (2002). Cortical involvement for action imitation of hand/arm postures versus finger configurations: An fMRI study. *NeuroReport*, *13*, 1599–1602.

Tokimura, H., Tokimura, Y., Oliviero, A., Asakura, T., & Rothwell, J. C. (1996). Speech-induced changes in corticospinal excitability. *Annals of Neurology*, *40*, 628–634.

Tomasello, M, & Call, J. (1997). *Primate Cognition*. Oxford: Oxford University Press.

Van Hoof, J. 1967. The facial displays of the catarrhine monkeys and apes. In D. Morris (Ed.), *Primate Ethology* (pp. 7–68). London: Weidenfield & Nicolson.

Visalberghi, E., & Fragaszy, D. (2001). Do monkeys ape? Ten years after. In K. Dautenhahn & C. Nehaniv (Eds.), *Imitation in Animals and Artifacts* (pp. 471–501). Cambridge, MA: MIT Press.

Vygotsky, L. S. (1934). *Thought and Language*. Cambridge, MA: MIT Press.

Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, *41*, 989–994.

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. *Neuron*, *40*, 655–664.

12 Organization of the Posterior Parietal Lobe and of Parietofrontal Connections

Giuseppe Luppino

The posterior parietal cortex (PPC) is generally thought to play a crucial role, in primates, in the analysis of high order sensory information, with special emphasis on visual information, which is then used for the guidance of motor behavior. One general issue, however, that is still a matter of controversy is whether the basic plans for the organization of PPC have been conserved in primates evolution and, in particular, in the evolution from monkeys to humans.

The main aim of this chapter is to address some crucial points of this controversy in the light of data, recently collected in our laboratory, on the organization of visual information processing in the monkey PPC.

The Issue of the Homology Between Monkey and Human PPC

In all primates, the intraparietal sulcus, a well-pronounced and constantly present sulcus, subdivides the PPC into a superior (SPL) and an inferior (IPL) parietal lobule. In spite of this macroscopical similarity, there is absolutely no definite consensus on whether the human and monkey PPC have a similar microstructural and functional organization.

From the microstructural point of view, basically, two markedly different positions have been proposed. One position is based on the observations of anatomists (Bonin & Bailey, 1947; Eidelberg & Galaburda, 1982), which favored a correspondence between the cytoarchitecture of the macaque and human PPC. Accordingly, in these studies the nomenclature firstly introduced by von Economo (1929) was adopted for referring, in both the macaque and human brain, to IPL and SPL architectonic areas with a similar location and considered to be homologous.

The other, more popular position is maintained on the basis of the studies of Brodmann (1909). According to Brodmann, the SPL and the IPL in the monkey consist principally of areas 5 and 7, respectively. In contrast, in the map of the human brain, which still presently represents the most widely used anatomical frame of reference also in the functional literature, areas 5 and 7 are both located in the SPL, whereas

the IPL is formed by two areas, 39 and 40, considered as newly generated areas particular to the human brain.

Although the Brodmann's position would imply, in evolutionary terms, an extremely fast change, with a complete architectonical reorganization of the parietal cortex in the evolution from monkeys to humans, the notion that the human IPL is a new evolutionary acquisition has been maintained also on the basis of functional considerations.

In humans, PPC lesions may produce a variety of syndromes, according to their location into either the SPL or the IPL. In particular, two of these syndromes, the optic ataxia and the neglect, are considered to represent paradigmatic examples of the disruptions of two distinct mechanisms of visual information processing in the human PPC. Optic ataxia is more frequently observed following lesions of the caudal SPL, and basically consists of impairment in the use of visual information for the organization of arm movements (Perenin & Vighetto, 1988). Neglect, which is usually observed following lesions of the caudal IPL, consists of a more profound, perceptual deficit, affecting the awareness of the space contralateral to the lesion (see, e.g., Vallar, 1998).

If Brodmann's scheme is accepted, then neglect, produced by a lesion of areas present only in the human brain, cannot find an explanation on the basis of monkey data. Consistent with this notion is considered one of the prevalent views of the way in which visual information is analyzed in the monkey PPC.

It has been well accepted since the early 1980s that in the monkey different aspects of visual information are analyzed in parallel, along two main cortical information processing pathways, an occipitoparietal one and an occipitotemporal one, generally referred to as the dorsal and the ventral visual stream, respectively. The original proposal of Ungerleider and Miskin (1982) was that the dorsal visual stream is mostly devoted to space perception, whereas the ventral visual stream is involved in visual recognition and discrimination of objects. Subsequent evidence, however, led to a reinterpretation of the role of these two visual streams, which basically derives from the observations of Milner and Goodale (1995). According to this view, the dorsal visual stream is involved in the analysis of visual information for action organization and, therefore, is also referred to as "vision for action" pathway. In contrast, visual perceptual processes are considered a province exclusively of the ventral visual stream. Accordingly, if the monkey PPC is not involved in perception, then no part of it may correspond to the human IPL.

A Modern View of Monkey PPC Organization

In recent years, anatomical and functional evidence was drastically changed the general view of the organization of the monkey PPC.

It is quite clear today that this cortical region consists of a multiplicity of distinct areas, involved in the analysis of specific aspects of sensory information and related to the control of specific effectors (figure 12.1A; for reviews, see Colby, 1998; Rizzolatti, Fogassi, & Gallese, 1997). Furthermore, it also well assessed that a large number of these areas are related to the analysis of visual information, alone or in combination with somatosensory information, and that these visually related parietal areas are located in both the SPL and the IPL, in contrast with the classical notion that the dorsal visual stream involves only the IPL.

PPC is linked by strong and reciprocal connections to the agranular frontal cortex, and classical anatomical studies showed that these parietofrontal connections are topographically organized, with the SPL and the IPL being mostly connected with more dorsal (PMd) and more ventral (PMv) parts of the premotor cortex. In recent years, we readdressed the issue of the topographic organization of these connections in the light not only of these data showing the existence of a multiplicity of parietal areas, but also of the discovery that the agranular frontal cortex is formed by several anatomically and functionally independent areas (figure 12.1A; for reviews, see Rizzolatti & Luppino, 2001; Rizzolatti, Luppino, & Matelli, 1998).

The main general finding of these studies was that the parietofrontal connections are organized in a rather specific way. Firstly, each agranular frontal area is target of a specific set of parietal areas. Furthermore, within each set of parietal areas projecting to any given agranular frontal area, one or two areas are the source of projections that are by far the strongest and defined as "predominant" projections, whereas the other areas are source of much weaker projections and defined as "additional" projections. Secondly, each parietal area, even if may project to more than one agranular frontal area, typically is source of predominant projections to a very restricted frontal sector. Thus, each parietal area appears to have a privileged target into the frontal cortex. If these privileged connections are taken into account, it is then possible to define, within the general framework of the parietofrontal connections, a series of largely segregated parietofrontal circuits, formed by parietal (most of them visually related) and frontal areas linked by predominant connections (figure 12.1B). Functional evidence, when available, indicates that parietal and frontal areas forming a given circuit have similar functional properties. This anatomical organization of the parietofrontal connections, then, represents the substrate for a parallel processing of different aspects of visuomotor transformations. It has been proposed that these parietofrontal circuits represent the functional units of the cortical motor system (Rizzolatti, Luppino, & Matelli, 1998).

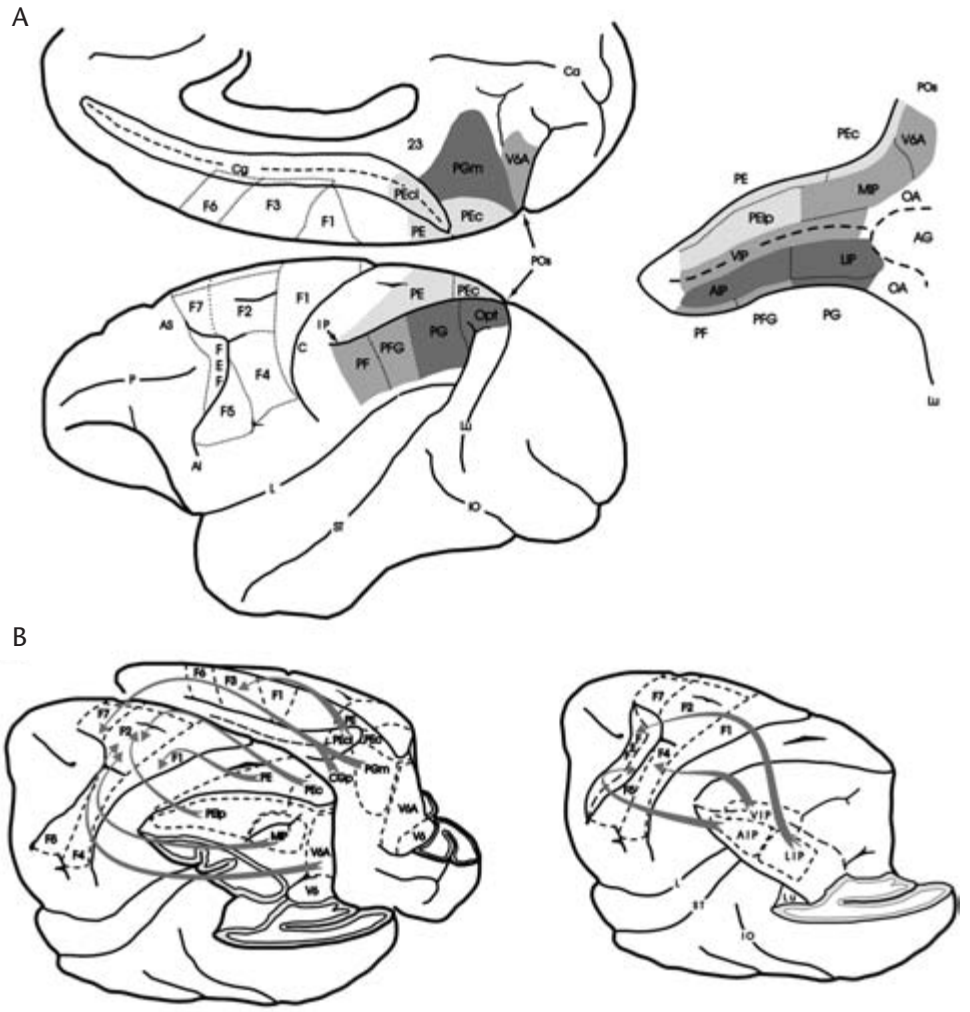


Figure 12.1

(*A*) Mesial and lateral views of the monkey brain showing the subdivision of PPC and of the agranular frontal cortex. The areas located within the intraparietal sulcus (IP) are shown in an unfolded view of the sulcus in the right part of the figure. For the definition of the posterior parietal and agranular frontal areas, see Rizzolatti et al. (1998). Purely visual, visual, and somatosensory and somatosensory areas are indicated in dark, intermediate, and light gray, respectively. (*B*) Summary view of the main projections from PPC to the agranular frontal cortex in the macaque monkey. (*Left*) Projections from the SPL; in this view of the brain the IPL and the occipital lobe have been removed, in order to expose the medial bank of the IP and the anterior wall of the POs. (*Right*) Projections from the lateral bank of the IP; in this view of the brain the IP was opened and the occipital lobe removed. AG, annectant gyrus; AI, inferior arcuate sulcus; AS, superior arcuate sulcus; C, central sulcus; Ca, calcarine fissure; Cg, cingulated sulcus; IO, inferior occipital sulcus; L, lateral fissure; Lu, lunate sulcus; P, principal sulcus; POs, parieto-occipital sulcus; ST, superior temporal sulcus.

A New Model of Visual Information Processing in the Monkey PPC

All together, the above-mentioned data show the existence of a multiplicity of largely segregated parietofrontal visuomotor circuits, related even to the same effector (e.g., the arm) and involving markedly different parietal (e.g., SPL vs. IPL) and frontal (e.g., PMd vs. PMv) sectors. These data, therefore, raise the question of which is their possible differential role in visual information processing and in motor control.

As it will be shown in this section, anatomical data recently collected in our laboratory, along with data from the literature, favor a new interpretation on the general organization of visual information processing in the macaque PPC. The hypothesis proposed in this chapter is that the various parietofrontal circuits, fed by the dorsal visual stream, can be grouped into two anatomically and functionally distinct main components, the dorso-dorsal (d-d) stream, formed by circuits involving the SPL and PMd areas and the ventrodorsal (v-d) stream, formed by circuits involving the IPL and, mostly, PMv areas and the FEF.

There are several arguments in favor of this subdivision of the dorsal visual stream into two major components, not only based on the differences in the parietal regions involved in the d-d and in the v-d streams (SPL vs. IPL, respectively) and in the markedly different roles in motor control played by their final target areas in the agranular frontal cortex (PMd vs. PMv and FEF). In particular, the major argument in favor of this subdivision is represented by the marked differences in the sources of visual information feeding the two streams.

The classical view on the anatomical organization of the dorsal stream is that this visual pathway is mostly under the influence of the so-called magnocellular pathway, which, at the cortical level, originates from layer IVB of V1. In this view, a nodal point of this pathway, crucial for the distribution of visual information to the PPC, is area MT/V5. This area, mainly through a relay into area MST, is the major source of visual inputs to the caudal IPL areas, considered to be final target into the PPC of this pathway. However, the discovery that also the caudal part of the SPL is involved in the analysis of visual information raised the question of the possible source of visual input to this parietal sector. It is quite clear today that another extrastriate area—V6—placed at the same level as MT/V5 in the hierarchy of visual extrastriate areas, represents a further main nodal point of the magnocellular pathway. As will be shown in the next subsections, V6 and MT/V5 project to markedly different parietal areas and represent the major source of visual information to the d-d and the v-d streams, respectively.

The D-D Stream

Area V6, located in the deepest part of the anterior wall of the parieto-occipital sulcus, represents the major source of visual input to the d-d stream.

Evidence provided in the early 1980s (Gattass, Sousa, & Cowey, 1985) showed that in the caudalmost part of the SPL, classically considered to be involved only in the analysis of somatosensory information, there is a visual area, termed PO, lying on the anterior wall of the parieto-occipital sulcus. These data were in full agreement with architectonic data showing that this cortical sector is all located within the limits of the occipital cytoarchitectonical domain (area 19 of Brodmann, 1909; area OA of von Bonin & Bailey, 1947, and Pandya & Seltzer, 1982). Later studies (Colby, Gattass, Olson, & Gross, 1988), on the basis of myeloarchitectonic data, confined area PO to about the ventral half of the wall and showed that this area is directly connected with V1 and other extrastriate areas.

The anterior wall of the POs was then the object of extensive electrophysiological studies by Galletti and collaborators (see Galletti et al., 2003), which led to the definition of two functionally distinct areas: a ventral and smaller area, defined as V6, corresponding only to the ventral part of area PO, and a more dorsal and larger area, defined as area V6A (figure 12.2A).

V6 is a retinotopically organized, purely visual area, with a complete representation of the contralateral visual hemifield, in which the representation of the central visual field is not emphasized as in other extrastriate areas. Virtually all V6 neurons are responsive to visual stimuli and their receptive fields, although larger than in the adjacent V3, are relatively small. In contrast, in V6A only about 60 percent of the neurons are responsive to visual stimuli. Their receptive fields are relatively large, and there is no clear topographic organization in the representation of the visual field. Nonvisual neurons in V6A are sensitive to gaze direction and/or saccadic eye movements, to somatosensory stimulation and to the execution of arm movements and represent a chief physiological criterion for demarcating the border of this area with V6.

These data, therefore, seriously challenged the notion that the whole anterior wall of the POs belongs to the occipital extrastriate cortex. For this reason, in a recent cytoarchitectonic study (Luppino, Matelli, Gamberini, & Galletti, 2003) we aimed to establish whether the two functionally defined areas V6 and V6A have also an anatomical counterpart and whether this cortical sector, as a whole or only in part, is located within the limits of Brodmann's area 19. The results showed that V6 and V6A correspond to distinct architectonic entities, which belong to different architectonical domains (figure 12.2B). In particular, on the basis of their general architectonic features, V6 was considered to belong to the occipital cytoarchitectonic domain and, therefore, to be located within the limits of Brodmann's area 19. In contrast, area V6A, formed by a ventral (V6Av) and a dorsal (V6Ad) architectonic subdivision, was considered to belong to the PPC.

Tracer injections restricted to V6 (Galletti et al., 2001) showed that this area is target of strong projections from layer IVB of V1 (figure 12.2C) and is connected with several

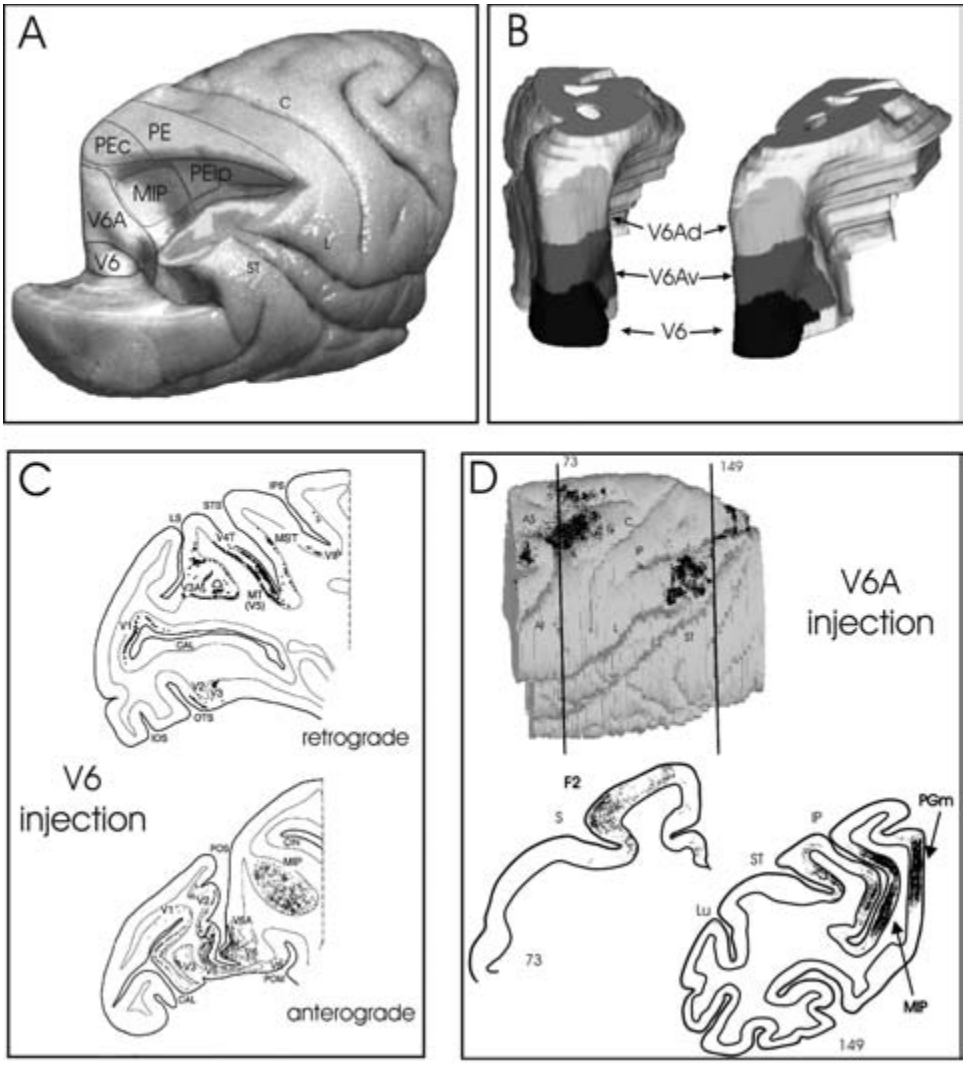


Figure 12.2

(*A*) Caudolateral view of a right hemisphere of the macaque brain in which the lateral bank and the occipital lobe were removed, in order to show the location of areas V6 and V6A. (*B*) Caudal (left) and caudolateral (right) views of the 3D reconstruction of the caudal part of a right SPL showing the location and extent of the cytoarchitectonic areas of the anterior wall of the POs. (*C*) Drawings of representative parasagittal sections showing the distribution of retrograde and anterograde labeling in the occipital and parietal cortices observed following injection of tracer (WGA-HRP) in area V6. (*D*) Distribution of labeled cortical neurons observed following injection of WGA-HRP in area V6A, shown in a dorsolateral view of the hemisphere and in two representative coronal sections taken at the levels marked by vertical lines. S, spur of the arcuate sulcus. Other abbreviations as in figure 12.1.

other occipital areas, including V2, V3, V3A, V4T and V5/MT. Connections with V5/MT are particularly strong and, according to the laminar distribution of the retrograde and anterograde labeling, were classified as lateral connections. V6 is not connected with areas of the frontal lobe and the connections with posterior parietal areas are almost completely limited to the caudalmost part of the SPL, in particular to V6A and MIP (figure 12.2C).

Thus, at the origin of the d-d and the v-d streams there are two areas, V6 and V5/MT, respectively, which are strongly interconnected each other, appear to be located at the same hierarchical level in the flow of visual information and are both target of layer IVB of V1, the cortical origin of the magnocellular pathway.

Area V6A, is not connected with other extrastriate areas than V6 and, in turn, is source of strong projections to PMd (Galletti et al., 2001; Marconi et al., 2001; Matelli, Govoni, Galletti, Kutz, & Luppino, 1998), in particular with the ventral and rostral part of area F2, which contains a representation of proximal and distal arm movements (figure 12.2D). V6A is also connected with IPL areas MST and 7a and particularly strong are the connections with two neighboring areas of the caudal SPL, MIP, and PGm. Therefore, connections attributed in previous studies to area PO with areas 7a and PGm (Colby, Gattass, Olson, & Gross, 1988) or with PMd (Tanné, Boussaoud, Boyerzeller, & Rouiller, 1995) are most likely due to the involvement of V6A by the tracer injections of these studies. MIP and PGm are also source of strong projections to different PMd sectors (Matelli et al., 1998). In particular, MIP mostly projects to ventrolateral F2, whereas PGm mostly project to the rostral PMd area F7.

All together these data strongly support the notion that V6 is an extrastriate area located at the origin of a relatively direct ("fast") visuomotor pathway (the d-d stream), which, in a few cortical steps, conveys visual information, with emphasis on the visual periphery, to caudal SPL areas (mainly V6A, but also MIP), the sources of parietofrontal projections to arm-related fields of PMd (figure 12.3). Therefore, the monkey caudal SPL very likely represents the homologue of the human caudal SPL, the cortical sector whose lesion produces optic ataxia.

The V-D Stream

MT/V5, mostly through the neighboring area MST, is the major source of visual input from the magnocellular pathway to the v-d stream. Well-known targets in the IPL of MT/V5 and MST are two areas located in the lateral bank of the intraparietal sulcus, the oculomotor area LIP and area VIP, and the caudal part of the IPL convexity, commonly referred to as area 7a (Boussaoud, Ungerleider, & Desimone, 1990). Area 7a represents the final target in the PPC of the dorsal visual stream in the original definition of Ungerleider and Miskin (1982) and is generally considered as a high order, purely visual area, in which retinal and extraretinal signals are integrated for visuospatial

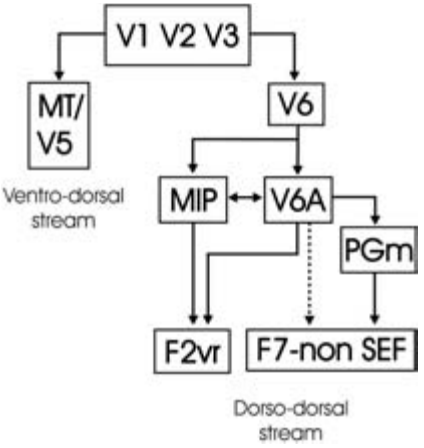


Figure 12.3
Schematic diagram of the dorso-dorsal visual stream. Arrowed lines indicate main connections, whereas the arrowed dashed line indicates a minor connection.

processing. Among the main connections attributed to this area are connections with areas LIP, PGm and prefrontal areas 8 and dorsal 46 (Andersen, Asanuma, Essick, & Siegel, 1990; Cavada & Goldman-Rakic, 1989).

According to the original architectonic definition of the Vogts (Vogt & Vogt, 1919), area 7a is a relatively large cortical area, which occupies about the caudal two thirds of the IPL convexity. Other architectonic studies, however, suggest a more complex subdivision of this cortical sector. In particular, Pandya and Seltzer (1982) defined in the IPL convexity, at least three cytoarchitectonic areas, a rostral, an intermediate and a caudal one, referred to as PF, PG, and Opt, respectively. Furthermore, a transitional sector was defined between PF and PF and referred to as PFG. If the two maps are compared each other, it is quite clear that area 7a of the Vogts corresponds to at least two areas, PG and Opt, according to the subdivision of Pandya and Seltzer. It is, however, common practice in the functional literature to refer to area 7a to also as area PG only. Independent evidence in favor of the view that the subdivision of the IPL convexity in two areas only (7a and 7b) is possibly too simplistic can be also found in the functional studies of Hyvärinen and colleagues (see Hyvärinen, 1981). According to these authors, in the IPL convexity (even within area 7a), there are different functional fields, distinguishable on the basis of the neuronal sensory properties and on the effectors to which they appear to be related.

In our laboratory, we recently addressed the issue of the anatomical organization of this cortical sector, the most important source of parietal afferents to the PMv, by combining cytoarchitectonics and tract tracing experiments. The aim was to define the

number and extent of the cytoarchitectonic areas forming the IPL convexity and their afferent and efferent connections.

Cytoarchitectonic data showed that the IPL convexity consists of four distinct areas, located at different rostrocaudal levels. By adopting the nomenclature of Pandya and Seltzer, these areas were referred to as PF, PFG, PG, and Opt. Thus, our data are in substantial agreement with that of Pandya and Seltzer, with the additional finding that we were able to identify area PFG as a distinct architectonic entity.

The cortical connections of the IPL areas were then assessed by placing tracer injections into each of the various identified architectonic areas of the IPL convexity and in the two adjacent areas located in lateral bank of the intraparietal sulcus, LIP, located more caudally and AIP located more rostrally.

One major finding of this hodological study is that the two caudal IPL areas, Opt and PG, which all together correspond to area 7a, have markedly different cortical connections. Area Opt (figure 12.4A), as expected, is the target of a strong input from area MST. Other strong connections of area Opt are with the adjacent area LIP, with the mesial parietal area PGm, and with the dorsal part of the dorsolateral prefrontal cortex (mainly dorsal area 46). Some connections were also observed with the agranular area F7. Area PG (figure 12.4B) is also target of strong projections from area MST, but lacks all the other connections observed for Opt. In contrast, area PG is connected with several arm-related fields of the agranular frontal cortex (mainly the rostral PMv area F5, but also the ventrorostral F2) and arm-related fields of the SPL (mainly areas MIP and V6A). These data, therefore, indicate that the connections commonly attributed to area 7a are observed only following injections in area Opt, whereas area PG, on the basis of its cortical connections, clearly appears to be an arm-related area.

A further main finding of this study is that the IPL territory target of MST projections extends beyond the limits of areas LIP and 7a. In fact, labeling into MST was observed also following injection in PFG, which, in turn, is the source of projections to the PMv areas F4 and F5 (figure 12.4C). Furthermore, in a case in which two different tracers were injected in the oculomotor area LIP and in the hand-related area AIP (figure 12.5, plate 17) labeling in MST was observed not only, as expected, following injections in LIP, but also following injection in AIP.

All together, these data indicate that in the IPL there is a set of distinct areas, all targets of area MST, but related to the control of different effectors and projecting in a differential way to the frontal lobe. All these areas can be considered to be located along the main pathway of the v-d stream.

A further finding of our study (but see also, e.g., Baizer, Ungerleider, & Desimone, 1991), which is very important for characterizing the v-d stream with respect to the d-d stream, was that all these MST-recipient areas of the IPL are also targets of afferents from the temporal cortex. These projections originate from two cortical sectors, the rostral two-thirds of the dorsal bank of the superior temporal sulcus (STS) and the

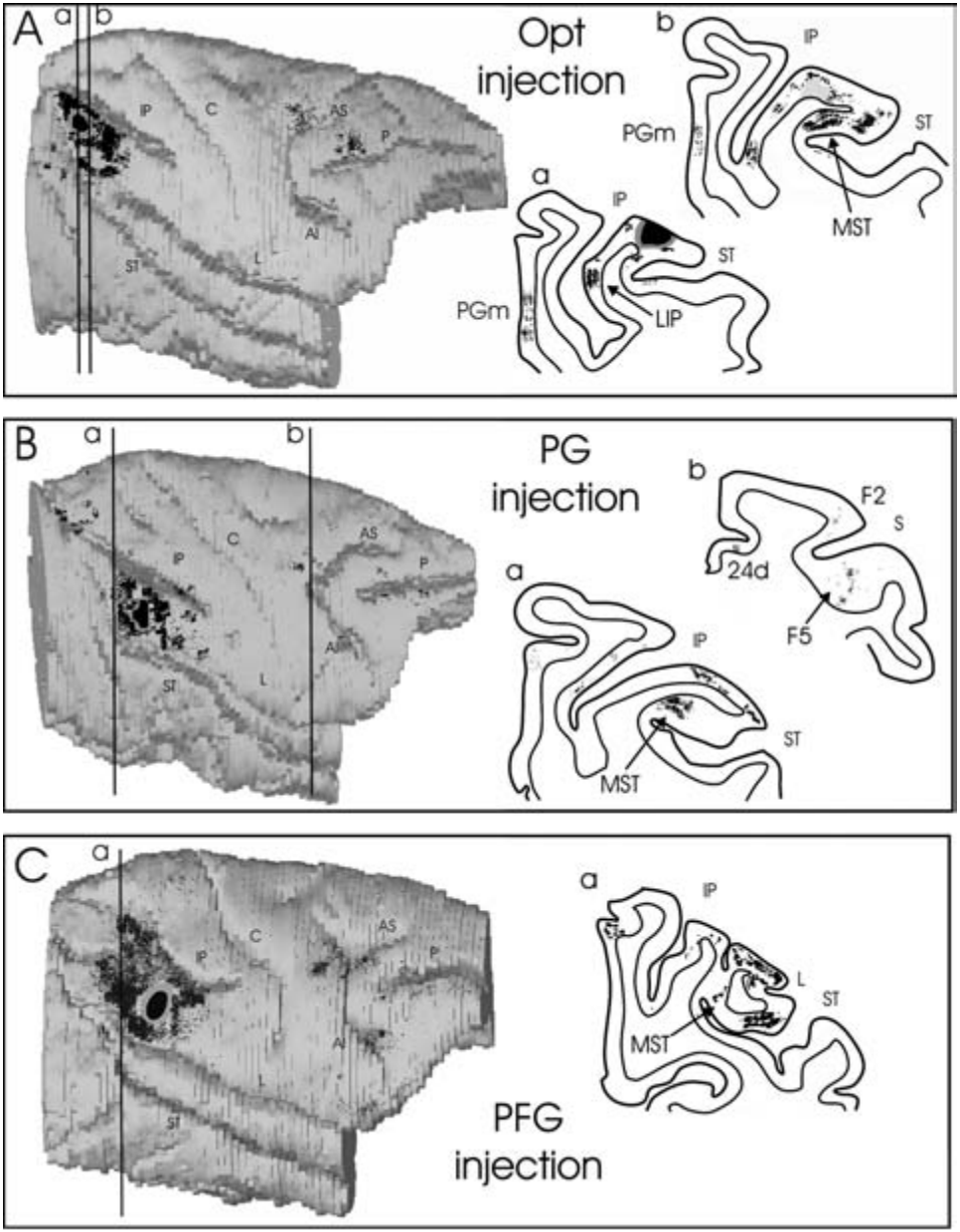


Figure 12.4

Distribution of labeled cortical neurons observed following WGA–HRP injections in areas Opt (*A*), PG (*B*), and PFG (*C*), shown in dorsolateral views of the hemispheres and in representative coronal sections taken at the levels indicated by vertical lines. Abbreviations as in figure 12.1.

inferotemporal cortex, both potential sources of different types of high-order visual information.

The rostral sector of the dorsal bank of the STS is generally referred to as the "superior polysensory area" (STP; see chapter 10, this volume). STP is a site of convergence of projections from somatosensory, auditory, and visual areas of both the dorsal and the ventral visual stream. Functional data are in line with this pattern of afferent projections showing that neurons in this area may have somatosensory, auditory, or visual receptive fields. Visually responsive neurons in STP may have very complex properties. All the various areas of the v-d stream, but area AIP, are targets of projections from STP (figures 12.5 and 12.6). Clusters of labeling were observed at different rostrocaudal levels of this cortical sector and, therefore appear to involve both the rostral (STPr) and the caudal (STPc) subdivision of STP.

All the various v-d stream areas are also targets of projections from areas more directly related to the ventral visual stream (figures 12.5 and 12.6). These projections originate from the fundal region of the STS (area IPa) or from the ventral bank of the STS, extending also on the convexity of the inferior temporal gyrus (area TE). Particularly strong were the projections to area AIP, which mostly originate from a large extent of area TE, an area located at the highest hierarchical level of the ventral visual stream. Neurons in this area have complex visual responses related to recognition and discrimination of object and, therefore, to the coding of object semantics (see chapter 13, this volume).

All together, these data indicate that the d-d and the v-d streams can be distinguished not only for the differences in the source of visual information from the magnocellular pathway, but also on the basis of afferents from the temporal cortex which target the v-d stream, but not the d-d stream (figure 12.7). It is important to note here that most of the various PPC areas (expecially those of the IPL) are also targets of different sets of afferents from other cortical sources, including the somatosensory, PPC, cingulate, and parahippocampal areas, which can further characterize the functional role of PPC areas. The analysis of these inputs is, however, outside the scope of this chapter, which is focused on the general organization of the flow of visual information through the monkey PPC.

Functional Considerations on the Temporal Projections to the V-D Stream

The projections from STP and from the ventral stream to the IPL are potential sources of high-order multisensory information, including the visual one, or of information related to object semantics. The functional role of these inputs in the processing of visual information along the v-d stream, still remain to be fully assessed. These projections, however, appear to be very helpful for explaining some of the functional properties recently described in some IPL areas.

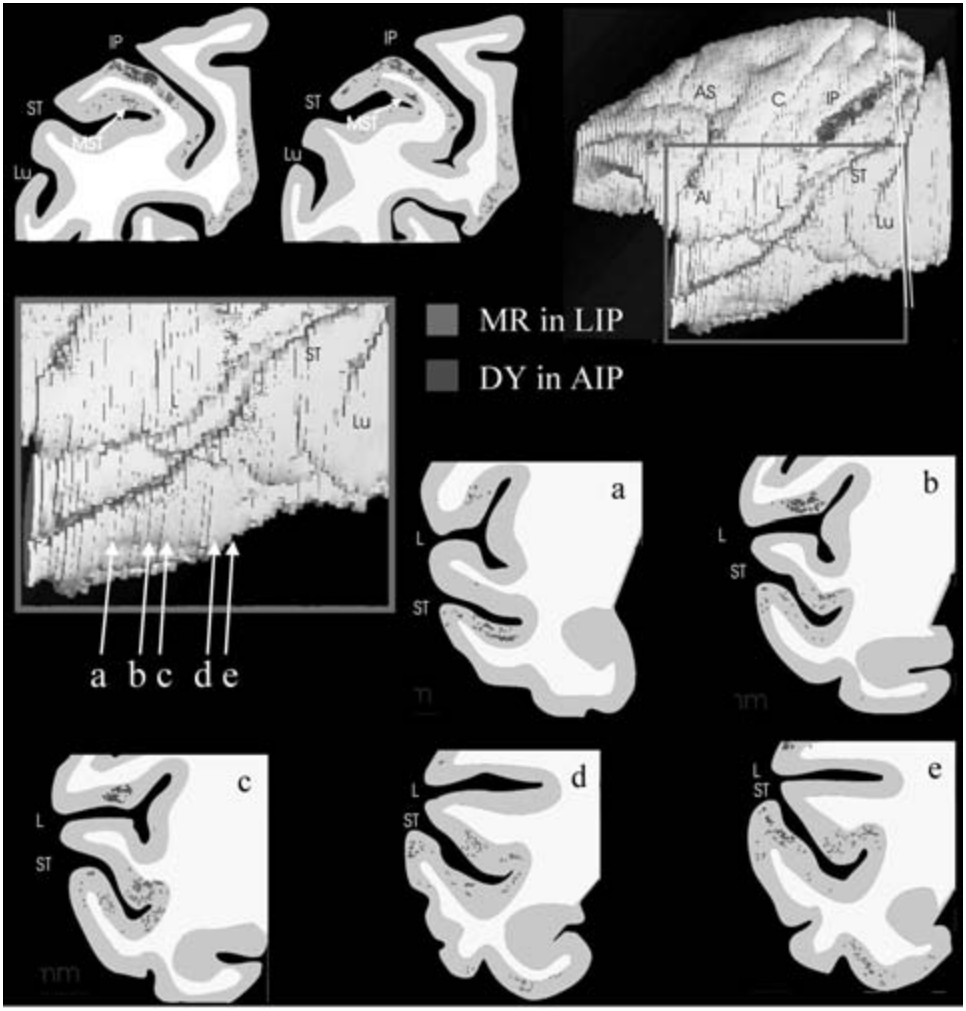


Figure 12.5
Distribution of labeled cortical neurons observed following fluorescent tracers injections in areas LIP (Diamidino Yellow, DY) and AIP (Microruby, MR). The upper part of the figure shows the distribution of the labeling on a dorsolateral view of the hemisphere (*right*) and in two representative coronal sections. The lower part of the figure shows the distribution of the labeling in a series of representative coronal sections through the temporal cortex, taken at the level indicated by vertical lines in an enlarged view of the hemisphere, centered on the temporal lobe. Abbreviations as in figure 12.1. See plate 17 for color version.

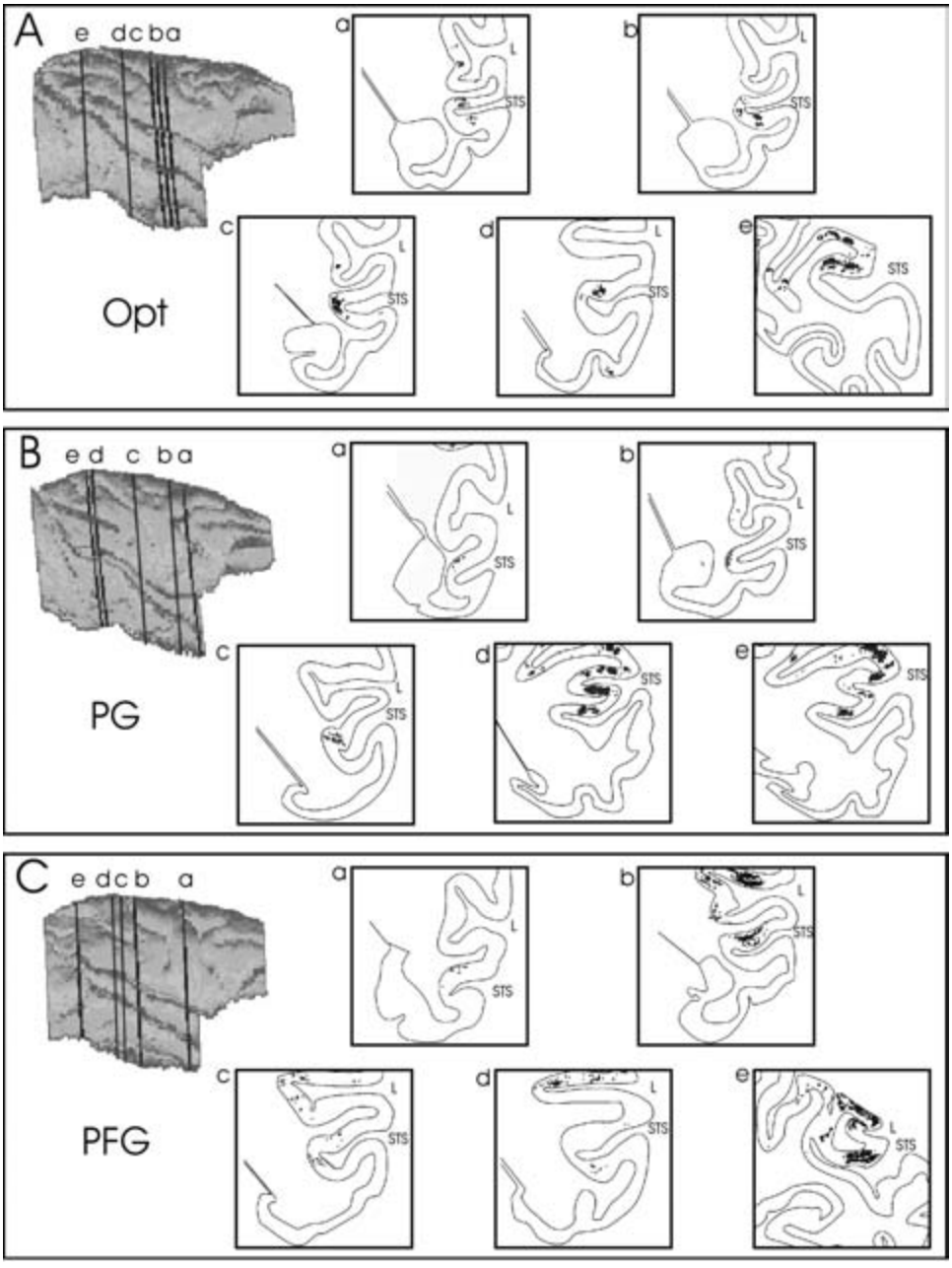


Figure 12.6

Distribution of labeled neurons in the temporal cortex observed following WGA–HRP injections (same cases as figure 12.4) in areas Opt (*A*), PG (*B*), and PFG (*C*), shown in representative coronal sections taken at the levels indicated by vertical lines on the lateral views of the hemispheres. Abbreviations as in figure 12.1.

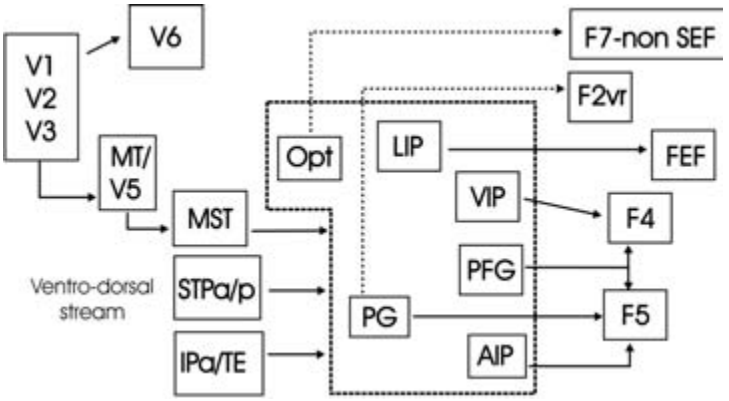


Figure 12.7
Schematic diagram of the ventrodorsal visual stream. The IPL areas of the ventrodorsal stream are included in the thick dashed line. Conventions as in figure 12.3.

For example, recent data showed that in the rostral part of the IPL (possibly in PFG) there are neurons with mirror properties (see chapter 11, this volume). It is possible that the complex visual properties of this class of neurons, related to the coding of actions made by other individuals, rely also on inputs from STP sectors in which neurons coding actions or biological motion have been described (chapter 10, this volume).

Furthermore, the strong projections from the inferotemporal cortex to area AIP, allow one to reexamine the well-established role of this area in visuomotor transformation for grasping. Area AIP, together with area F5, is involved in a highly selective parietofrontal circuit, which plays a crucial role in those neural processes that, from the visual analysis of the object intrinsic properties (e.g., size, shape, orientation), lead to the choice of the most appropriate patterns of distal movements for object grasping and manipulation (see Rizzolatti & Luppino, 2001). It has been proposed that these processes rely, basically, on a "pragmatic" analysis of the object intrinsic properties, which occurs along an occipitoparietal pathway that passes from the extrastriate cortex through a caudal and ventral part of the lateral bank of the intraparietal sulcus, reaching area AIP (Sakata, Taira, Murata, & Mine, 1995). The data presented in this chapter strongly suggest that the "pragmatic" processing of this visuomotor circuit relies also on information on object semantics.

In general, it is largely accepted that the ventral stream must interacts at some cortical level with the motor system, in order to allow a control of motor behavior based on perception. This interaction is generally considered to occur in the prefrontal cortex through connections between areas selective targets of the ventral stream (e.g., 45 and 12) and areas connected with the premotor cortex or the PPC (e.g., 46). The

significant projections from the ventral stream to the IPL represent, however, a more direct route that allows the access of semantics to the motor system. Thus, in the monkey IPL, there is a group of areas in which visual information from both the dorsal and the ventral visual stream is integrated with information on motor programs. It is likely that it is this type of integration that is at the basis of perceptual processes, and, therefore, the v-d stream could represent the visual pathway crucial for control of motor behavior, based on perceptual processes.

The general subdivision of the monkey PPC proposed in this chapter is based on the analysis of the general organization of the flow of visual information through this cortical region. It is important to note here that, for example, the v-d stream includes a group of several areas, each of them with distinct connections with other sensory and nonsensory areas. Furthermore, each of these areas has distinct patterns of projections to the frontal motor areas and, therefore, will be presumably involved in specific aspects of motor control.

Concluding Remarks

Data reviewed in this chapter indicate that visual information processing in the monkey PPC follows two main pathways (the d-d and the v-d streams), which run through different parietal lobules (SPL and IPL, respectively) and appear to have a differential role in the control of motor behavior.

These data, therefore, suggest that in the evolution from monkey to humans the basic plans for the organization of PPC may have been conserved and are consistent with the scheme in which the intraparietal sulcus divides PPC in a similar way in monkeys and humans. It should be noted, however, that in the monkey all the various IPL areas appear to be involved in the control of specific effectors. It is then possible that the monkey IPL functionally represents the homolog of the human left IPL circuits subserving spatial functions, whereas the human right IPL acquired, in the evolutionary processes leading to lateralization of functions, high-order spatial representations more detached from action.

Acknowledgments

The results presented here were obtained thanks to grant from EU (Contract QLRT-2001-00746) and MIUR.

References

Andersen, R. A., Asanuma, C., Essick, G., & Siegel, R. M. (1990). Corticocortical connections of anatomically and physiologically defined subdivisions within the inferior parietal lobule. *Journal of Comparative Neurology*, *296*, 65–113.

Baizer, J. S., Ungerleider, L. G., & Desimone, R. (1991). Organization of visual inputs to the inferior temporal and posterior parietal cortex in macaques. *Journal of Neuroscience*, *11*, 168–190.

Bonin, G. von, & Bailey, P. (1947). *The Neocortex of Macaca mulatta*. Urbana: University of Illinois Press.

Boussaoud, D., Ungerleider, L., & Desimone, R. (1990). Pathways for motion analysis: Cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. *Journal of Comparative Neurology*, *296*, 462–495.

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Groshirnrinde*. Leipzig: Barth.

Cavada, C., & Goldman-Rakic, P. S. (1989). Posterior parietal cortex in rhesus monkey: I. Parcellation of areas based on distinctive limbic and sensory corticocortical connections. *Journal of Comparative Neurology*, *287*, 393–421.

Colby, C. L. (1998). Action-oriented spatial reference frames in cortex. *Neuron*, *20*, 15–24.

Colby, C. L., Gattass, R., Olson, C. R., & Gross, C. G. (1988). Topographical organization of cortical afferents to extrastriate visual area PO in the macaque: a dual tracer study. *Journal of Comparative Neurology*, *269*, 392–413.

Economo, C. von (1929). *The Cytoarchitectonics of the Human Cerebral Cortex*. London: Oxford University Press.

Eidelberg, D., & Galaburda, A. M. (1982). Symmetry and asymmetry in the human posterior thalamus. I. Cytoarchitectonic analysis in normal persons. *Archives of Neurology*, *39*, 325–332.

Galletti, C., Gamberini, M., Kutz, D. F., Fattori, P., Luppino, G., & Matelli, M. (2001). The cortical connections of area V6: An occipito-parietal network processing visual information. *European Journal of Neuroscience*, *13*, 1572–1588.

Galletti, C., Kutz, D. F., Gamberini, M., Breveglieri, R., & Fattori P. (2003). Role of the medial parieto-occipital cortex in the control of reaching and grasping movements. *Experimental Brain Research*, *153*, 158–170.

Gattass, R., Sousa, A. P. B., & Cowey, E. (1985). Cortical visual areas of the macaque: possible substrates for pattern recognition mechanisms. In C. Ghagas, R. Gattass, & C. G. Gross (Eds.), *Pattern Recognition Mechanisms* (pp. 1–20). Vatican City: Pontifical Academy of Sciences.

Hyvärinen, J. (1981). Regional distribution of functions in parietal association area 7 of the monkey. *Brain Research*, *206*, 287–303.

Luppino, G. , Matelli, M., Gamberini, M., & Galletti, C. (2003). Cytoarchitectonic organization of the anterior bank of the parieto-occipital sulcus in macaque brain. *Society for Neuroscience Abstracts*, 701.19.

Marconi, B., Genovesio, A., Battaglia Mayer, A., Ferraina, S., Squatrito, S., Molinari, M., Lacquaniti, F., & Caminiti, R. (2001). Eye-hand coordination during reaching. I. Anatomical relationships between parietal and frontal cortex. *Cerebral Cortex*, *11*, 513–527.

Matelli, M., Govoni, P., Galletti, C., Kutz, D. F., & Luppino, G. (1998). Superior area 6 afferents from the superior parietal lobule in the macaque monkey. *Journal of Comparative Neurology, 402*, 327–352.

Milner, A. D., & Goodale, M. A. (1995). *The Visual Brain in Action*. Oxford: Oxford University Press.

Pandya, D. N., & Seltzer, B. (1982). Intrinsic connections and architectonics of posterior parietal cortex in the rhesus monkey. *Journal of Comparative Neurology*, *204*, 196–210.

Perenin, M.-T., & Vighetto, A. (1988). Optic ataxia: A specific disruption in visuomotor mechanisms. *Brain*, *111*, 643–674.

Rizzolatti, G., Fogassi, L., & Gallese, V. (1997). Parietal cortex: From sight to action. *Current Opinion in Neurobiology*, *7*, 562–567.

Rizzolatti, G., & Luppino, G. (2001). The cortical motor system. *Neuron*, *31*, 889–901.

Rizzolatti, G., Luppino, G., & Matelli, M. (1998). The organization of the cortical motor system: New concepts. *Electroencephalography and Clinical Neurophysiology*, *106*, 283–296.

Sakata, H., Taira, M., Murata, A., & Mine, S. (1995). Neural mechanisms of visual guidance of hand action in the parietal cortex of the monkey. *Cerebral Cortex*, *5*, 429–438.

Tanné, J., Boussaoud, D., Boyerzeller, N., & Rouiller, E. M. (1995). Direct visual pathways for reaching movements in the macaque monkey. *NeuroReport*, *7*, 267–272.

Ungerleider, L., & Mishkin, M. (1982). Two visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of Visual Behavior* (pp. 549–586). Cambridge, MA: Mit Press.

Vallar, G. (1998). Spatial hemineglect in humans. *Trends in Cognitive Neuroscience*, *2*, 87–97.

Vogt, O., & Vogt, C. (1919). Ergebnisse unserer Hirnforschung. *Journal für Psychologie und Neurologie*, *25*, 277–462.

13 A Prototype of *Homo faber*: A Silent Precursor of Human Intelligence in the Tool-Using Monkey Brain

Atsushi Iriki

The Tool-Use Gifted Human Brain: An Intellectual Primitive

Tool use has long been believed to be a peculiar characteristic of human intellectual ability, as we define ourselves as *Homo faber*, meaning "man who uses and makes tools." Some two million years ago, the first stone flakes used as "tools" appeared in Africa. They were "Oldowan pebble tools" manufactured by *Homo habilis*, a protohominid whose name means "skillful man." This primitive tool seems not more than a direct extension of the strength of the human knuckle and fist, perhaps used to crack something very hard. Since then, varieties of more and more sophisticated tools evolved along with hominid evolutionary processes (Wynn, 1996). Eventually, when diversity of tools suddenly "exploded" about 30,000 years ago, tools were no longer merely a direct extension of our bodily functions. That is, they became complexly structured based on causal relationships of various parts comprising the tool as a whole. And today, the hominid instinct of tool invention has resulted in the blossoming of our modern technologies.

Initiation of bipedal locomotion has been emphasized as a critical factor for developing the abilities of using tools—the freed forelimbs (the hands) gave the human brain the extra capacity of resources to acquire the fine dexterity that is an essential requisite for tool use (Paillard, 1993). But, skillfulness per se should not necessarily constitute a condition sufficient for the emergence and development of tool use. Which capacity of brain faculty, and what kind of mental process, was the major driving force, perhaps triggered by bipedalism, for hominids to induce the rapid development of use and manufacture of tools and technology?

Now, consider our introspection when using a most simple type of tool, for example, a stick grasped by hand. By holding a stick in the hand, we can extend our reaching distance. Then, by this "extended" hand, we can feel with the tip, hit with the tip, and retrieve something using the tip of the tool. Hence, we can (1) physically, (2) perceptually, and (3) functionally extend our innate body structure. Further, a great advantage of having a tool at hand is not just extending our body parts in space in a

fixed manner—it can be grasped at any time we want and can be abandoned at any time we do ***not*** want to use it. Thus, one can select at any time any adequate tool that exactly matches the given situation. If this faculty is further advanced, one may become able to plan to use, and even manufacture the tool to exactly fulfill one's purpose in one's mind, including secondary tools to make an intended tool, and this process would eventually lead to a modern technology by their combinations.

In conclusion, the principal advantage of the tool as an extension of our innate body is that it does not constrain various aspects of the image of the body, and we can modulate and choose it as we like at any given moment according to the current intention—we acquired a "freedom" to configure our functional body structure according to our free will. Therefore, flexible modification of internal representations free from physical constraints of the actual physical world should be the key element of the expression of tool-using abilities. These abilities would be extrapolated into higher intellectual cognitive abilities of hominids, such as language and metaphysical thoughts, which might coevolve with tool use and technology.

Tool Use in Nonhuman Primates

Classical work on tool use in nonhuman primates was, perhaps, first described by Kohler (1927) in chimpanzees; seeing a banana hanging high from the ceiling, a chimpanzee "thought" for a while, and then suddenly climbed up some boxes, jumped on top with a stick at hand, and took the banana. This report was the first evidence that described the "insightful" and "spontaneous" use of tools by a nonhuman species. This, then, opened the research field of the use of tools in various other animal species.

In primates lower than apes, evidence for tool use has been rather fragmentary (Tomasello & Call, 1997). Even in apes other than chimpanzees, tool-use behavior has not been as frequently observed. Macaque monkeys seldom use tools in their wild habitat, and spontaneous usages upon demand are ambiguous. In Japanese macaques, occasional observation of spontaneous tool-use behavior is reported, but evidence for "insightful" tool use has not been presented. On the other hand, in several different species of New World monkeys, evidence of spontaneous tool use has been repeatedly reported.

These discrepancies regarding the emergence of tool use in different primate species have given rise to a debate on the strict definition of "usage" of tools, and also the definition of "tool" per se, trying to determine the mental and/or neural substrates that would be the essential factors for the expression of tool-using behaviors. Various questions were raised. Are there critical differences in the machineries of the brain between the primates that use tools and those that do not? Or, should tool use be made possible to be observed in lower primates when properly trained, even though they do not express those abilities in naïve and wild atmospheres? Then, would there

be any unique and discontinuous concrete neural correlates that subserve highly intelligent tool use and technology in humans, or, alternatively, are they merely continuous and quantitative differences from other primates?

In this chapter, I will first demonstrate that various forms of higher cognitive functions, previously thought to be particular to humans, could be induced by training in Japanese macaques. Then their neural correlates will be identified, both physiologically and molecular-genetically. Finally, based on these lines of evidence, I would like to propose a hypothesis regarding how these silent precursors of the monkey brain could be recruited to be expressed during the course of evolution to finally represent human intelligence with which we define ourselves today.

Neural Precursors of Intelligence in the "Reinforced" Monkey Brain

Monkeys Using Tools

Although Japanese macaques in their natural habitat rarely use tools spontaneously (Tomasello & Call, 1997), we could train them to retrieve distant objects beyond the limit of reach of their innate arm, using a rake as a tool (Iriki, Tanaka, & Iwamura, 1996). This learning process usually takes two weeks to accomplish, and could never be shortened (Ishibashi, Hihara, & Iriki, 2000); this evidence has been repeatedly confirmed so far in more than thirty monkeys without exception. (The behavioral characteristics during this training process are described in a later section.) When a food pellet was dispensed beyond reach, the monkey wielded the tool and pulled the food closer so as to retrieve it with the other hand (figure 13.1A). Thus, the reachable distance, beyond which the monkeys ignored food pellets, became greater when the tool was available.

Psychological evidence indicates that when a tool is used as an extension of the hand, it is incorporated into the putative body image, or body schemata—the knowledge about the dimensions, posture and movement of the hand in the environmental space—perhaps formed in the parietal cortex by integration of intrinsic (somatosensory) and extrinsic (visual) information related to the corresponding body parts (Head & Holmes, 1911). Indeed, human parietal lesions result in neglect or extinction of the existing body parts, and conversely, amputees could perceive the subjective experience of nonexisting body parts (the "phantom limb" phenomenon) perhaps having this sort of mental representation previously acquired (Berlucchi & Aglioti, 1997). Based on the above considerations, experiments were performed to attempt to determine the neural correlates of the above putative body schema, or body images, in the tool-using monkey parietal cortices (Iriki, Tanaka, & Iwamura, 1996).

Chronic recordings were made from neurons in the bank of the intraparietal sulcus, where hierarchically processed somatosensory information (Iwamura, 1998) adjoins

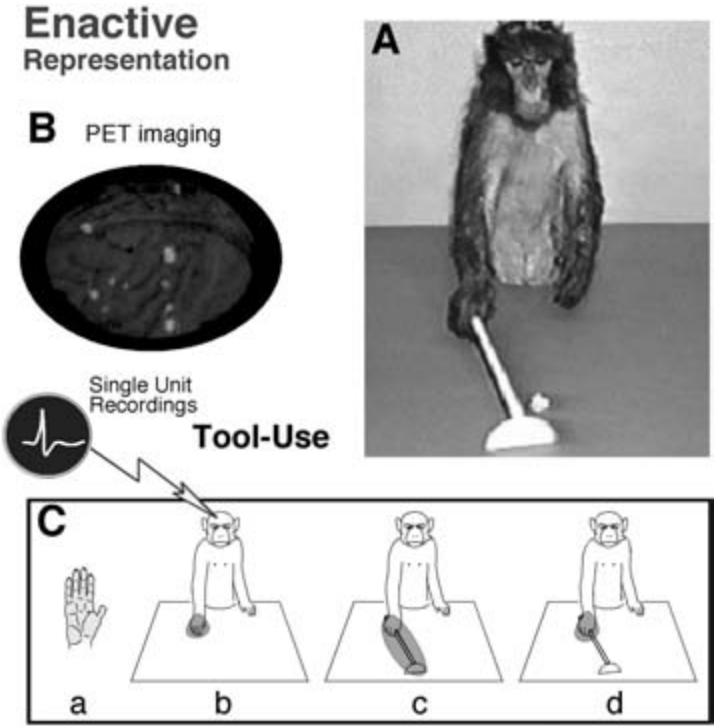


Figure 13.1

Neural correlates of assimilation of the hand-held tool into the "enactive representation" (Bruner, 1966) of the hand, or the "body schema" (Head & Holmes, 1911). (*A*) Monkeys were trained to use a rake to retrieve food (a piece of an apple) placed beyond the reach of the innate hand. (*B*) PET (positron emission tomography) imaging of brain activation during tool use, obtained by subtracting global brain activation during manipulating a pseudo-tool stick from that during retrieving food with the rake. (*C*) Changes in bimodal receptive field properties of a representative neuron upon tool use: (a) somatosensory receptive field (RF) (b–d) visual RFs before (b) and immediately after (c) tool use, and that when just passively holding the rake (d). (*A* and *C* modified from Maravita & Iriki, 2003; *B* adapted from Obayashi et al., 2001.)

the information on spatial vision processed along the dorsal stream (Ungerleider & Mishkin, 1982). Neurons in this cortical area (contralateral to the hand using tools), responding to both somatosensory and visual stimulations, namely "bimodal" neurons, were analyzed. The somatosensory receptive field (RF) was identified by a handheld probe, a tiny paintbrush, passive manipulation of joints or active hand use. Visual RF was defined as a territory of the space in which visual stimuli evoked action potentials. The most effective form of visual stimulus for inducing neural activity was that moving towards the somatosensory RF.

A large number of intraparietal bimodal neurons appeared to code the image of the hand by integrating somatosensory and visual information—these neurons have visual RFs that encompass, and are locked onto, their somatosensory RFs located at the hand and forearm. Before tool use, the somatosensory RFs of the representative neurons shown in figure 13.1 was located on the glabrous skin of the palm and digits (figure 13.1Ca), whereas the visual RFs before tool use encompassed the hand where the somatosensory RFs were located (figure 13.1Cb).

After the monkey held the rake and repeated food-retrieving actions for five minutes, RFs were reexamined. In more than half of the bimodal neurons, visual RFs enlarged to become elongated along the axis of the tool to include its entire length (figure 13.1Cc), as if the image of the tool had been incorporated into that of the hand, while the somatosensory RFs remained unchanged. After the monkey retrieved food without using the tool for three minutes, the once-expanded visual RFs diminished even if the monkey kept holding the tool during the recording period (figure 13.1Cd). Neurons exhibiting elongation of the visual RFs along the rake had the somatosensory RFs in the hand/forearm. Thus, we call all these neurons the "distal type."

On the other hand, neurons with somatosensory RFs located at proximal body parts (including upper arm, shoulder, neck and face, namely the "proximal type") had visual RFs that cover the space accessible with the innate arm. After tool use, these neurons became responsive to stimuli presented in a wider region, namely, within the space accessible to the handheld rake.

All of these use-dependent expansions occurred only when the monkeys held a tool and intended to use it as an extension of their hands—expansion was not induced when monkeys did not intend to use a tool, even though they held it. Because physical properties of both the appearance of the hand holding the rake and the movements of the scanning probe in space were identical between the two situations, the difference in visual RFs should reflect the monkey's introspections, whether the monkey was "looking upon" the rake as an extension of his hand or not.

These findings may constitute the neural correlates for modification of the body image comprising bases of introspective assimilation of the tool into our own body. Indeed, these neurons were found most heavily in the arm/hand region of the postcentral gyrus but not in the digit region, perhaps because the rake is an extension of

the hand and forearm but not of the digits in the present experimental situation. Thus, the above neural evidence found in tool-using monkeys might represent neural substrates existing in common with humans to subserve introspective assimilation of tools into our body schemata.

While scanning the entire space (this typically takes two to three minutes) to identify visual RFs in the above described neurophysiological experiments, monkeys were asked to keep their hand quiet on the table—neurons discharging during tool-using actions were discarded from the data. Because of the restricted experimental conditions, we could conclude that the responses should purely represent the perceptual domain of how the monkey regarded the introspective body image, and not be related to any intention of producing tool-using actions. On the other hand, however, also because of this technical limitation, it has been difficult to identify whether monkeys are actually utilizing this intraparietal machinery during tool use by manipulating the body image as its dynamic property. This limitation was finally overcome by the use of PET (positron emission tomography) imaging of the monkey brain during tool use (Obayashi et al., 2001). By subtracting global brain activation during the manipulation of a pseudo-tool stick from that during the retrieval of food by using the rake, the intraparietal region contralateral to the hand using the tool (where the above-described bimodal neurons reside) together with some other motor-related brain areas were confirmed to be activated during this task (figure 13.1B).

Monkeys Playing Video Games

When playing a video game or using teleoperator systems, we feel our self-image projected onto the video monitor as an extension of ourselves. The general belief until now was that monkeys, and even apes, were not capable of doing this but were just able to "utilize" the image on the monitor to guide bodily movements (Matsuzawa, 2001; Tomasello & Call, 1997). However, the above conclusions were reached from behavioral analyses per se without any empirical measure, nor any verbal report of course, about the introspective self in nonhuman primates, which has been believed impossible to obtain.

Now, such a self-image is examined if it can be coded by the same group of bimodal neurons in the monkey intraparietal cortex, as described above, coding modified body images upon tool use (Iriki et al., 2001). In this experiment, in which monkeys were trained to use the tool under video-captured images projected on a monitor (figure 13.2A), the visual RFs of the bimodal neurons were formed in the video screen-images on the screen scanned by artificial probe superimposed by chromakey effects. Thus, the image of the hand on the monitor (figure 13.2Ba), eventually extended along the tool upon its usages, should be coded by these bimodal neurons as an extension of the hand belonging to the monkey's own body. This would constitute the first empirical demonstration that the introspective body image of the animals is actually

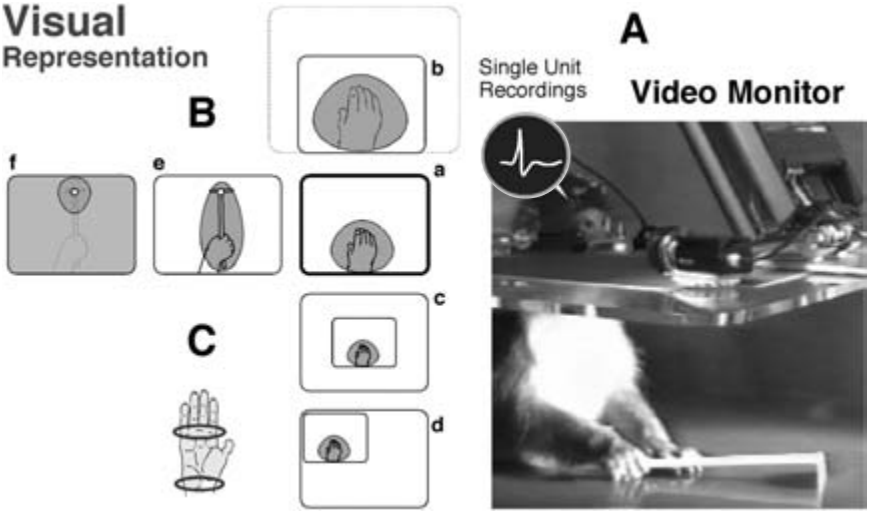


Figure 13.2
Neural correlates of the perception of self-image projected onto a video monitor, representing the "visual representation" (Bruner, 1966) of the hand. (*A*) Monkeys, prevented from direct vision by an opaque plate below the eyes, were trained to retrieve food by using tools while relying only on the visual feedback provided through video captured images projected on a monitor. (*B*) Visual receptive field (RF) properties of a representative neuron, as assessed with an artificial visual probe superimposed on the image by chromakey effect. Visual RFs formed around the hand (a) were modified according to expansion (b), compression (c), or change of position (d) of the visual image on the monitor. After tool use, visual RFs extended along the tool (e), and were formed around the tool-tip as a functional extension of the hand (f) when only the spot at the tip of the tool was left visible by luminancekey effect. (*C*) Somatosensory RF of this neuron. (Modified from Maravita & Iriki, 2003.)

projected on the video monitor to recognize as an extension of "self." Having this mechanism in the brain, we could feel "reality" in virtual apparatus or video games.

Further, the size and position of visual RFs of the bimodal neurons were modified according to expansion (figure 13.2Bb), compression (figure 13.2Bc), or change of position (figure 13.2Bd) of the visual image on the video monitor. During the scanning, posture and position of the hand (and of course its size) were not actually altered during modification of the screen image. Thus, these properties of the visual RFs would represent neural correlates for the intentionally controllable visual representation of their own body. In addition, the neuron presented here responded not only to the natural image of the hand on the monitor, but also to a sign that functionally substitutes for the actual hand in the monitor. That is, when only the spot at the tip of the tool left visible using luminancekey effects, the visual RF once extended along the

tool (figure 13.2Be) was now formed around the spot (resembling a cursor on a computer monitor) as a functional extension of the hand (figure 13.2Bf). These observations may suggest an evolutionary precursor for introspective manipulation of an abstract sign existing in the monkey brain—a symbolic representation of own body might be furnished, although not functioning in natural and naive conditions but ready to be expressed by proper training, or by "education."

Thinking Monkeys

We can perform purposeful hand actions even in the dark blindly, perhaps by relying on images of the hand in the mind. Activities of the same group of bimodal neurons in the monkey intraparietal cortex as described above were shown to represent such subjective images of the body, which are not actually visible but created only in the mind (Obayashi, Tanaka, & Iriki, 2000).

Initially confirmed was that visual stimuli moving into the space encompassing their somatosensory RFs in the hand activated these neurons to form the visual RFs coding the hand-image. Then, after the hand was made invisible by covering with an opaque plate (a liquid crystal shutter that becomes opaque within 100 msec), the visual RFs were reexamined and were found to persist over the plate immediately above the invisible somatosensory RFs. When the hand was moved invisibly under the plate, the visual RFs moved over the plate to follow the invisible somatosensory RFs. Hence, monkeys were shown able to maintain and update (perhaps relying on proximal proprioceptive information) the subjective body-image in the mind, and such images were coded by intraparietal bimodal neurons. Further, the same group of neurons was found recently to code not only the location of the body parts, but also the internal images of the "movements" of the hand in space by combining information about directions of joint displacements and moving visual stimuli in space (Tanaka et al., 2004).

The above lines of evidence suggest that long-lasting introspective representation of the body parts and their movements in space exists in this brain area, and it is maintained and updated by referring, perhaps, to somatosensory information arising from the proximal part of the arm. This would represent neural correlates of "visually induced somesthetic imagery" or "somesthetically updated visual image" of own body parts. Thus, subjective representation of such body images should reside in the intraparietal area of the monkey cerebral cortex. Indeed, phenomena directly relevant to this have been shown to exist in human parietal cortex (see Blakemore & Sirigu, 2003, for review). This would constitute a form of an image free from physical constraint of the actual external world. Having this form of representation, which could be manipulated and combined in the mind, we should be able to imagine complex bodily actions, before execution, to achieve the intended goal most efficiently, and perhaps eventually acquire perception of causal relationships between elementary actions.

Thus, this might constitute precursors of basic neural mechanisms possibly subserving hominid higher cognitive functions, which will be described in later sections.

Genetic Recruitment of Evolutionary Precursors

Gene Expression by Tool-Use Training

Japanese macaques could use a rake as a tool to retrieve distant pieces of food when they were "properly" trained under specific laboratory conditions (Ishibashi, Hihara, & Iriki, 2000). This learning process usually took two weeks to complete, as follows: (1) During the first one week or so, the monkeys merely moved the rake around randomly or just pulled it straight without purpose, appearing as if they did not understand that the rake could be used as a tool to extend their reach (figure 13.3Aa). (2) Then, by about the tenth to twelfth day, they started to behave as if they had "discovered" or realized the function of the rake, and tried to manipulate it. During this period, manipulation of the rake was divided into two phases, first moving it sideward and placing its tip beyond the location of the food, then pulling it closer to retrieve the food (figure 13.3Ab). During this period, the success rate to retrieve the food increased dramatically (figure 13.3B). Thus, the "cognitive learning" process was considered to be completed by this time, although the movement of the tool still remained rather slow and awkward. (3) Finally, by not later than two weeks after the initiation of training, tool-use movement became smooth, rapid and efficient, which would indicate that "motor learning" was completed (figure 13.3C).

The above processes required an entire two-week period, never less, leading us to suppose that some sort of molecular genetic processes must be taking place in the brain to reorganize neural circuitry to express a novel faculty—coding body images modified by voluntary intention. During this period, when the success rate rapidly increased (figure 13.3B, arrow), induction of the expression of immediate-early-genes (figure 13.3D) (Ishibashi, Hihara, Takahashi, & Iriki, 1999) and brain-derived neurotrophic factor, its receptor *trk*B, and NT-3 (Ishibashi et al., 2002a, 2002b) were revealed selectively in the area where bimodal neurons coded the image of the hand into which the tool was incorporated, that is to say, among the anterior bank of the intraparietal sulcus immediately posterior to the postcentral hand/forearm representations. After the whole learning processes were completed, this augmented expression compared with the control level was never again detected.

Assuming that some sort of "evolutionary pressure" has reinforced the primate brain to express tool-using abilities, precursors (or basic blueprint) of the human tool-using brain faculties should exist in the lower primate brain waiting for adequate environmental demand to push forward those silently hidden machineries. During the "cognitive learning" period described above, demanding training under experimental conditions might have mimicked such pressure, enabling us to observe neural

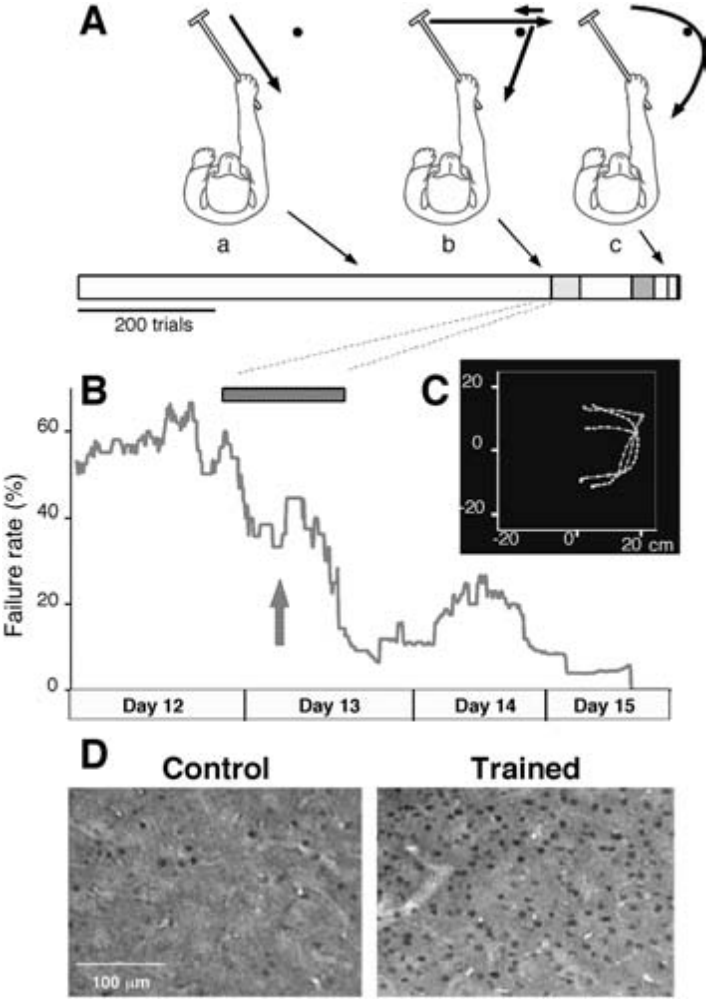


Figure 13.3

Development of the skill accompanied by gene expression during monkey tool-use learning. (*A*) Behavioral characteristics of tool usage, which consisted of just pulling straight unpurposefully during the first week (a), and then becoming composed of two phases (b), first sideward and then pulling movements, finally becoming smooth, rapid and efficient (c). (*B*) Learning process, illustrated by transition of failure rate (ordinate) between the twelfth and fifteenth day after initiation of training, when "cognitive learning" seemed to be achieved dramatically (see text). (*C*) Changes in trajectories of tool-tip movements during the "motor learning" process. Different colors represent movements exhibited during the period shown by the corresponding colors in (*A*). Each dot represents 33 msec. (*D*) Augmented expressions of an immediate-early-gene protein (Zif 268) in the anterior bank of the intraparietal sulcus contralateral to the hand using the tool (trained side) compared with the homotypical area of the other hemisphere (control). Dark spots represent Zif 268 positive cells revealed by immunohistochemistry. (*A–C* modified from Ishibashi, Hihara, & Iriki, 2000.)

correlates of experimentally "induced" functions such as tool use. Thus, this paradigm, using functions not expressed in the wild, happened to be very efficient for scientific studies, because the experimenter could control the degree of expressions in the laboratory for empirical analyses; such controlled quantifications would be difficult if expressed "spontaneously" uncontrolled in the natural habitat.

Development of Novel Neural Circuitry by Tool-Use Training

Given the above evidence for genetic expression in the intraparietal cortex during tool-use learning, the next logical question to be answered is which concrete neuronal processes can be induced by those genetic mechanisms. After tool-use training, it was repeatedly shown, as explained above, that a group of bimodal neurons coded variously modified body images depending on tasks. However, such bimodal neurons with distinct visual responses could hardly be seen in naive or untrained monkeys. In other words, visual responses are much less in the naive monkey brain, but become evident by training. Thus, visual input to these areas could have been reorganized to increase strong enough to code altered body images by integration with somatosenory inputs.

When retrograde tracer was injected into the anterior bank of the intraparietal sulcus of the trained monkeys, apparent labelings that do not exist in control animals were detected in cortical areas including the temporoparietal junction (TPJ) (Hihara et al., 2002). Hence, we hypothesized that the TPJ area might be a source of visual information that was augmented in the intraparietal cortex by tool-use training. Then, in turn, an anterograde tracer (BDA, biotinylated dextran amine) was injected into the TPJ area of tool-use-trained and naive adult Japanese monkeys, and the anterogradely labeled terminal arborization of single axons in the intraparietal cortex was compared (Hihara, Notoya, Tanaka, Ichinose, & Iriki, 2003). In control animals, anterogradely labeled axons were found only at the fundus of the intraparietal sulcus and a few boutons only among deeper cytoarchitectonic layers, while small numbers of labeled fibers were detected in the white matter at the shallower portion of the anterior bank of the intraparietal sulcus. In the trained monkeys, however, widely divergent terminal fields with complex arborizations were found most typically among superficial layers II and III throughout the entire bank of the intraparietal sulcus, with a large number of boutons scattered among these layers. Electron microscopic observations detected BDA-positive asymmetric synaptic terminals, with a large active zone and round vesicles, making contacts with dendritic spines of postsynaptic neurons.

This augmentation of corticocortical projections suggested that a novel mode of somatosensory-visual integration is developed morphologically, as a concrete modification of neural circuitry by training, in order to organize adequate manipulation of the body image, which is essential for using tools as an extension of the hand.

Linking Monkey Brain and Human Brain Functions

Development and Evolution of Internal Representations

The presently demonstrated neural correlates of the various aspects of the manipulated body images would provide clues for the developmental as well as evolutionary pictures of higher cognitive functions. During the earliest stage of human postnatal development, a child's field of view is restricted to personal and immediately adjacent peripersonal space. The body image of this period should be a type of sensorimotor intelligence, which is unconsciously acquired through experiencing various actions in the environmental space as accustomed patterns of action, thus constituting an "enactive representation" (Bruner, Olver, & Greenfield, 1966). The "body schema," postulated in 1911 by Head and Holmes, and its neural correlates in the monkey parietal cortex we showed earlier in this chapter, would delineate this subconscious aspect of the body image. This stage of the body image should contribute, through direct intracortical communications between parietal and motor cortices, to automated planning and execution of most efficient body actions in the space. Hence, this sort of representation should constitute the requisite for handling the most primitive "Oldowan artifacts," which were merely a direct extension of the function of innate hand, as described earlier in this chapter.

As children become familiar with the surrounding space, they achieve (by the age of 9 to 10) abilities to handle action-free visual images of their own body. Arriving at this stage, they become conscious of body image and admit the strong component of manipulation as a necessary aid to imagery. These "visual representations" or "ikonic representation" (Bruner, Olver, & Greenfield, 1966) have a perceptual domain that, when combined with actual input, would recall nonexisting subjective representation. And vice versa, cognitive and imaginary subjective experiences recall the engram of perceptual sensory experience. Thus, our way of visual representation of the world becomes free from action, which dissociates a spatially organized internal schema from supporting actions. The results illustrated that the neural responses coding self on the monitor are the demonstration of neural correlates of such a visual body image at this stage, indicating that macaque monkeys (along the course of primate evolution, like human children during development) attained an intentionally controllable representation of their own body.

When the above-described representation was further advanced, it would become totally free from physical constraints of the actual world to become a symbolic one (Bruner, Olver, & Greenfield, 1966). Indeed, restricted damage of this cortical area in human patients could result in "asymbolia," dysfunction of the symbolic processes. Whether or not macaque monkeys can exhibit symbolic processes is still an open question. However, the currently demonstrated objective observation of the subjective

mind, free from actual visual input, would represent evolutionary precursors of such "pre-symbolic" abilities.

This evidence may be linked to the dramatic explosion of hominid higher cognitive functions, which were first exhibited by the development and abundance of various kinds of stone tools some 30,000 years ago, eventually leading to the evolution of human language or to metaphysical thoughts, the neural mechanisms of which are still to be uncovered. In the following sections of this chapter, I will discuss some experimental evidence that could represent such aspects of higher cognitive functions, possibly linking functions between monkey brain and human brain.

Precursor of Mechanistic Technology?

Studies on primate tool-use behavior not only recorded the usage of tools but also examined the extent to which animals understood functional meanings of the tool, or the causal relationship between an intended tool-use action and the results obtained through it (Tomasello & Call, 1997; Matsuzawa, 2001; Povinelli, 2000). They came to conclude that, although neural mechanisms for object manipulation with tools might be similar, a huge difference exists between human and nonhuman repertoire of tool use—only humans have an insight into causality-based structures among objects (including tools) in the external world. Thus, perception of causality could constitute fundamentals of modern technologies (Wolpert, 2003).

As for recent experiments with monkeys, they were required to combine two different types of tools rationally so as to retrieve a reward (Hihara, Obayashi, Tanaka, & Iriki, 2003). Food was placed at a distance, reachable only with a long rake, but the monkeys could not reach the long rake directly. Instead, they could reach a short rake, but with it they could not reach the food (figure 13.4Aa, plate 18). In this situation, monkeys could easily solve the problem by using the short rake to get the long rake (figure 13.4Ab), exchanging tools (figure 13.4Acd), and get the food (figure 13.4Ae). This behavior was accomplished after a very short period of trials within a single day (figure 13.4B), in remarkable contrast to the initial basic training of using the tool, which took at least two weeks (figure 13.3B).

The above results indicate that once basic skill was learned through extensive training, perhaps accomplished through molecular genetic processes that allow reorganization of neural circuitry, the application of once-acquired basic skill could be accomplished rather easily, perhaps subserved by different brain mechanisms. The latter mechanism was studied by monkey PET imaging using a similar but slightly different tool-using skill (Obayashi et al., 2002). In that task, a food pellet was delivered into transparent tubing and the monkey had to initially push it using a rake to roll it out of the tubing, and then retrieve the pellet by pulling it in with another rake. The brain activation pattern obtained by subtraction of single tool use from this applied

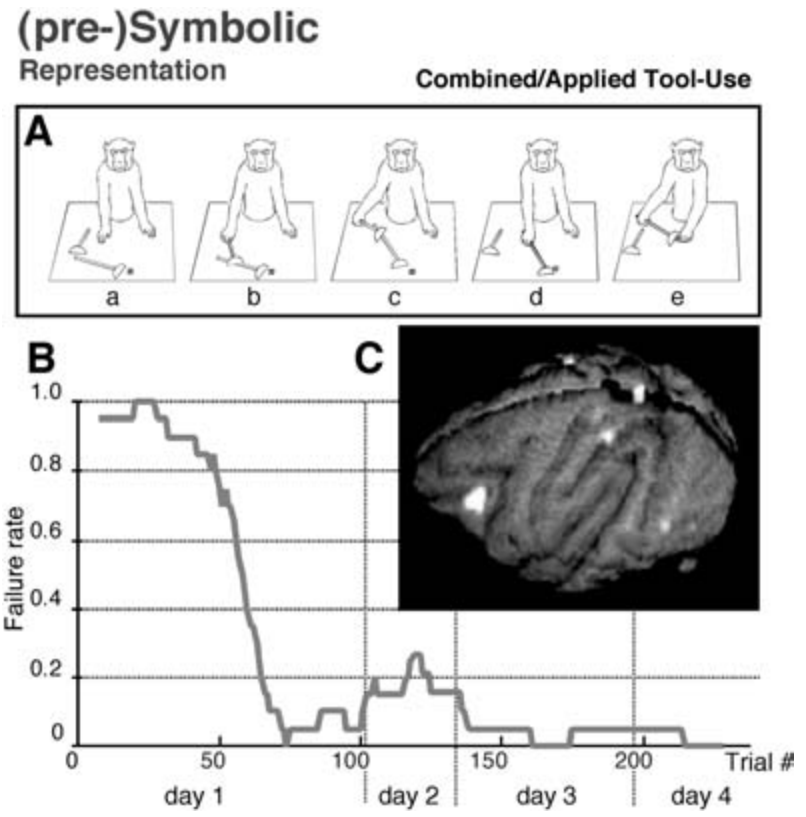


Figure 13.4

Monkeys' behavior to combine two different types of tools to retrieve the reward, perhaps employing "(pre-)symbolic representation" of the body image (Bruner, 1966). (*A*) The food was placed at a distance only reachable by long rake, but monkeys could not reach the long rake directly; instead, they could reach a short rake, but with which they could not reach the food (a). Monkeys took the long rake by using the short rake (b), exchanged the tools (cd) and took the food (e). (*B*) Learning process of this behavior was completed after a very short period of trials within one day. (*C*) PET imaging of brain activation pattern obtained by subtraction of single-tool-use from applied tool-use task; that is, the food pellet was delivered into the transparent tubing and the monkey had to initially push it using a rake to roll it out of the tubing, and then retrieve it by pulling it in, using another rake. (*A* and *B* modified from Hihara, Yamada, Iriki, & Okanoya, 2003a; C adapted from Obayashi et al., 2002.) See plate 18 for color version.

tool-use task indicates the presence of prefrontal activity in addition to that in the intraparietal area where basic learning takes place (figure 13.4C).

This suggests that prefronto-intraparietal interaction is essential for this applied usage of tools. The mechanism could be further extended to the perception of the causal relationships of the mode of elementary tool usages, through intentionally controlled manipulation of the internally created body images. Further, this would lead to development of an "insight" or perception of the causal structure of events, and intelligent usages of mechanistic tools in humans (Goldenberg & Hagmann, 1998). Thus, these would represent evolutionary precursors of the seeds of modern technology beyond simple tools (Wolpert, 2003).

Unlike simple "tools," "machine" has is own internal state and self-generated motion properties, constituting basic elements of modern technology. The above results could be extrapolated to brain mechanisms handling these "advanced tools" (Obayashi et al., 2004), which have never been determined before. Such mechanisms may possibly resemble those of social interactions with other human and/or nonhuman individuals. Thus, the above aspects could be further extended toward basic knowledge of humanoid robotics and, even more, to future technological social environments.

Precursor of "Naming"?

Vocal production and its usage in nonhuman primates share common features with primitive human language (Seyfarth, 1987). Primate vocalizations may designate objects or events in the external world or represent the affective state. However, previous studies focused on their ontogeny, or on their classification in restricted social settings (Hauser, 1996). Here we have observed that Japanese monkeys being trained to use tools spontaneously differentiated their coo-calls, a common social vocalization of wild monkeys, to ask for either the food or the tool (Hihara, Yamada, Iriki, & Okanoya, 2003). This would represent an evolutionary precursor of "naming."

In this study, monkeys were initially trained to use a rake to retrieve a reward, and to produce coo-calls asking for the reward. During the process of tool training, these monkeys showed a degree of plasticity in their coo-calls associated with the tool context. One of the monkeys was trained for the tool-use task, but he spontaneously emitted coo-calls when requesting a tool to obtain a reward. These calls differed from a regular food coo-call in their acoustic compositions. For the other monkey, three conditions of tool-assisted reward retrieval were attempted: (1) food was presented out of reach when the monkey produced the first call, and then (2) a tool was presented out of reach upon the second call, or (3) food was set out of reach and a tool was presented within reach when the monkey produced the first call. These calls emitted under three different conditions differed in their acoustic parameters and differentiation proceeded as the session advanced. It must be emphasized that we *never*

explicitly trained the monkeys to differentiate calls, and that they spontaneously emitted distinctive calls under different tool-use conditions. Vocalizations of nonhuman primates have been believed to be inherited and unmodifiable during a lifetime. However, we observed that monkeys spontaneously altered a repertoire of calls, suggesting an emergence of a novel vocal repertoire and referential use of the altered calls.

We speculated about the neural mechanisms subserving the above phenomenon in the following manner. Activation of the cingulate-midbrain pathway has been shown to produce primate vocalization, but this pathway lacks the control of higher cerebral cortices (Jurgens, 1990). Our observations indicate that monkeys can regulate calls voluntarily, and this requires a higher cognitive process. Tool-use behavior in primates is often discussed in relation with syntactical manipulations of treelike structures, which would represent "syntactic" aspects of language evolution (Matsuzawa, 2001). In contrast, here we observed a spontaneous differentiation of coo-calls and their referential use during tool-associated behavior. This process might involve a gradual change of emotional vocalizations into intentionally controlled vocalizations by associating them with consciously planned tool use. Tool use might induce referential use of vocal signals that could reinforce the cortical control of the midbrain vocal center. Tool use, thus, could also be an origin of "semantic" aspects, or the emergence of the perception of meaning of language. This would shed light on the origin of the "naming" behavior, which has not yet been well considered.

Conclusion

In this chapter it is described that, although Japanese macaques rarely use tools spontaneously in the wild, they could be trained to use a rake to retrieve distant food. During the learning phase, augmented expressions of immediate-early-genes, neurotrophic factors and receptors are involved. By these training-induced genetic expressions, reorganization of the neural circuitry was "induced" to develop a novel mode of somatosensory-visual integrations, thereby enabling this brain area to code modified body image upon tool use. Further, Japanese monkeys could express, by training, various higher cognitive abilities (body images on the video monitor, insightful tool use, and so on) that were previously believed not to exist in these lower primates.

Given the above evidence, we could imagine that a precursor (or basic blueprint) of human tool-use abilities was already furnished in our monkey ancestor brain, and had been pushed to full expression by some sort of "evolutionary pressure." Demanding training of tool use might have mimicked such a pressure to activate some silent neurogenetic mechanisms. This idea would give rise to the following general questions concerning the study of primate higher cognitive functions:

1. Which functions are expressed naturally and spontaneously in the wild habitat?
2. Which functions could not be expressed even with reinforcement?
3. Which functions could be "induced" by proper training, or education, and, in turn, which methods are adequate to induce unexpressed functions?

Thus, the tentative answer to the initial questions posed at the beginning of this chapter would be: The basic machinery furnished in the brain seems continuous, and differences should be a quantitative but not qualitative ones, among those primates that use or not use tools, and the only differences existed according to whether they were expressed or not expressed depending on a triggering factor in an environmental demand in the living habitat.

In the cases described in this chapter, the necessity to use tools might have triggered the higher cognitive functions to be induced. In the hominid evolutionary processes, the emergence of bipedal locomotion together with the initiation of the use of Oldowan artifacts might have served as the triggering factors. Although basic factors ready to subserve the faculty could be continuous, there could be possible differences between species that exhibit tool-using functions or not. Such an additional factor would be the one controlling the threshold level of spontaneous expressions of these functions, which wait to be uncovered. This scheme would not only provide clues to developmental as well as evolutionary aspects of existing higher cognitive functions of primates, but would also open a novel research platform from which to pose questions about what potential faculties future humans might acquire.

References

Berlucchi, G., & Aglioti, S. (1997). The body in the brain: Neural bases of corporeal awareness. *Trends in Neuroscience*, *20*(12), 560–564.

Blakemore, S. J., & Sirigu, A. (2003). Action prediction in the cerebellum and in the parietal lobe. *Experimental brain Research*, *153*(2), 239–245.

Bruner, J. S., Olver, R. R., & Greenfield, P. M. (1966). *Studies in Cognitive Growth*. New York: Wiley.

Goldenberg, G., & Hagmann, S. (1998). Tool use and mechanical problem solving in apraxia. *Neuropsychologia*, *36*(7), 581–589.

Hauser, M. D. (1996). *The Evolution of Communication*. Cambridge, MA: MIT Press.

Head, H., & Holmes, G. (1911). Sensory disturbances from cerebral lesions. *Brain*, *34*, 102–254.

Hihara, S., Notoya, T., Tanaka, M., Ichinose, S., & Iriki, A. (2003). Sprouting of terminal arborization of the temporoparietal afferents by tool-use learning in adult monkeys. *Society for Neuroscience Abstracts*, Program no. 939.15. On line.

Hihara, S., Obayashi, S., Tanaka, M., & Iriki, A. (2003). Rapid learning of sequential tool use by macaque monkeys. *Physiology & behavior*, *78*(3), 427–434.

Hihara, S., Tanaka, M., Notoya, T., Ojima, H., Ichinose, S., & Iriki, A. (2002). Cortical connections of monkey intraparietal cortex representing modified body-images during tool use. *Society for Neuroscience Abstracts*, Program no. 282.19. On line.

Hihara, S., Yamada, H., Iriki, A., & Okanoya, K. (2003). Spontaneous vocal differentiation of coo-calls for tools and food in Japanese monkeys. *Neuroscience Research*, *45*(4), 383–389.

Iriki, A., Tanaka, M., & Iwamura, Y. (1996). Coding of modified body schema during tool use by macaque postcentral neurones. *Neurorepotr*, *7*(14), 2325–2330.

Iriki, A., Tanaka, M., Obayashi, S., & Iwamura, Y. (2001). Self-images in the video monitor coded by monkey intraparietal neurons. *Neuroscience Research*, *40*(2), 163–173.

Ishibashi, H., Hihara, S., & Iriki, A. (2000). Acquisition and development of monkey tool-use: behavioral and kinematic analyses. *Canadian Journal of Physiology and Pharmacology*, *78*(11), 958–066.

Ishibashi, H., Hihara, S., Takahashi, M., Heike, T., Yokota, T., & Iriki, A. (2002a). Tool-use learning selectively induces expression of brain-derived neurotrophic factor, its receptor trkB, and neurotrophin 3 in the intraparietal multisensory cortex of monkeys. *Cognitive Brain research*, *14*(1), 3–9.

Ishibashi, H., Hihara, S., Takahashi, M., Heike, T., Yokota, T., & Iriki, A. (2002b). Tool-use learning induces BDNF expression in a selective portion of monkey anterior parietal cortex. *Molecular Brain Research*, *102*(1–2), 110–112.

Ishibashi, H., Hihara, S., Takahashi, M., & Iriki, A. (1999). Immediate-early-gene expression by the training of tool-use in the monkey intraparietal cortex. *Society for Neuroscience Abstracts*, *25*, 889.

Iwamura, Y. (1998). Hierarchical somatosensory processing. *Current Opinions in Neurobiology*, *8*(4), 522–528.

Jurgens, U. (1990). *Neurobiology of Comparative Cognition.* Hillsdale, NJ: Erlbaum.

Kohler, W. (1927). *The Mentality of Apes*. New York: Humanities Press.

Maravita, A., & Iriki, A. (2003). Tools for the body (Schema). *Trends in Cognitive Science* (in press).

Matsuzawa, T., ed. (2001). *Primate Origins of Human Cognition and Behaviour.* Berlin: Springer-Verlag.

Obayashi, S., Suhara, T., Kawabe, K., Okauchi, T., Maeda, J., Akine, Y., Onoe, H., & Iriki, A. (2001). Functional brainmapping of monkey tool use. *NeuroImage*, *14*(4), 853–861.

Obayashi, S., Suhara, T., Nagai, Y., Maeda, J., Hihara, S., & Iriki, A. (2002). Macaque prefrontal activity associated with extensive tool use. *NeuroReport*, *13*(17), 2349–2354.

Obayashi, S., Suhara, T., Nagai, Y., Okauchi, T., Maeda, J., & Iriki, A. (2004). Monkey brain areas underlying remote-controlled operation. *European Journal of Neuroscience*, *19*(5), 1397–1407.

Obayashi, S., Tanaka, M., & Iriki, A. (2000). Subjective image of invisible hand coded by monkey intraparietal neurons. *NeuroReport*, *11*(16), 3499–3505.

Paillard, J. (1993). The hand and the tool: The functional architecture of human technical skills. In A. Berthelet & J. Chavaillon (Eds.), *The Use of Tools by Human and Non-Human Primates* (pp. 36–46). Oxford: Oxford University Press.

Povinelli, D. (2000). *Folk Physics for Apes*. New York: Oxford University Press.

Seyfarth, R. M. (1987). *Primate Societies*. Chicago: University of Chicago Press.

Tanaka, M., Obayashi, S., Yokochi, H., Hihara, S., Kumashiro, M., Iwamura, Y., & Iriki, A. (2004). Intraparietal bimodal neurons delineating extrinsic space through intrinsic actions. *Psychologia*, *47*(2), 43–78.

Tomasello, M., and Call, J. (1997). *Primate Cognition*. New York: Oxford University Press.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), Analysis of Visual Behavior (pp. 549–586). Cambridge, MA: MIT Press.

Wolpert, L. (2003). Causal belief and the origins of technology. *Philosophical Transactions of the Royal Society, London A: Mathematics, Physics, Engineering, and Sciences*, *361*(1809), 1709–1719.

Wynn, T. (1996). The evolution of tools and symbolic behaviour. In A. Lock & and C. Peters (Eds.), *Handbook of Human Symbolic Evolution* (pp. 263–287). Oxford: Oxford University Press.

14 Parietal Mechanism of Selective Attention in Monkeys and Humans

Claire Wardak, Suliann Ben Hamed, and Jean-René Duhamel

Together with the frontal and temporal cortices, the parietal cortex is sometimes referred to as an associative cortex. The term *associative* indicates that it is a region of massive convergence of information coming from all over the brain. Its lesion in humans leads to altered cognitive functions in the representation of space, in the representation of one's body, in the preparation of goal directed movements toward it and in the orientation of attention toward salient objects in that space. This has led to the general idea that this cortical region is implicated in the representation of space, in the sensorimotor transformations required for goal-directed movements and in selective attention processes. G. Luppino discusses in this same volume the specific parietofrontal cortical projections and how such a network contributes to the construction of goal-directed movements. On a different line, A. Iriki discusses how the properties of specific parietal neurons participate in the construction of a body schema and, more specifically, integrate tool use in this schema.

For our part, we will focus on the contribution of parietal cortex to selective attention mechanisms both in humans and in monkeys. A large network subtends attentional processes, including parietal, frontal (mainly the frontal eye field) and subcortical (mainly superior colliculus and pulvinar) regions. We are still far from understanding the contribution of each of these structures to this complex cognitive process. What follows is an attempt to compare the role of different parietal regions in different aspects of attentional processing in both species and to discuss the possible homologies between these different parietal regions in humans and monkeys.

One of the first electrophysiological reports of attentional modulations in primate parietal cortex is that of Mounstcastle and colleagues, who show that attentive fixation enhances the excitability of neurons in the inferior parietal lobule as compared to a free gaze condition (Mountcastle et al., 1981). A recent study confirms these pioneering observations and demonstrates that the modulation of neuronal excitability is due to complex differences in the shape of the receptive fields during the two conditions (Ben Hamed et al., 2002). It shows that the hot spot of the receptive field can shift toward the fovea during fixation, and that enhancement of the neuron's response

can be limited to the classical receptive field or extend beyond it. The overall effect of these modulations is to increase the visual resources around the fovea during fixation and distribute them across the visual field during free gaze.

Attentional modulations have since been described in several extrastriate, parietal, prefrontal, and subcortical regions. We will focus on parietal correlates of attention processes. Early studies following that of Mountcastle et al. describe two types of attentional modulations: spatially specific modulations and baseline modulations, each of which may play a specific role in attentional functions. Interestingly, as will be seen here, these fundamental neuronal signatures of attentional processes in the parietal lobe have recently been observed using neuroimagery techniques in humans.

Spatially Specific Attentional Modulations

The introduction of behavioral tasks requiring the engagement of attention in a particular peripheral spatial location led to one of the first electrophysiological demonstrations of a modulation of the visual neuronal signals by selective spatial attention in the parietal cortex of monkeys. The animals were required to fixate a central target and to detect a change in luminosity of a peripheral target located inside the receptive field of the visual neuron being recorded. More than half of the parietal neurons (mainly located in areas 7a and LIP—lateral intraparietal area; see figure 14.1) show a modulation of their visual response when the monkey is attending the location of the stimulus (Bushnell et al., 1981). Three-quarters of these neurons show an enhancement of their response, the remaining quarter showing a suppression. This attentional modulation of the visual response is specific to the spatial locus of attention, and is present whenever the monkey has to use information provided by the stimulus to guide its behavior, should it be a manual response or a saccade (Colby et al., 1996). This amplification of the neuronal response has been considered as an electrophysiological marker of spatial allocation of attention. However, attentional modulations can also take the form of a suppression. In 7a, for example, more than half of the neurons show a suppressed response to the reappearance of a visual stimulus at an attended location, while less than 5 percent show an enhancement (Steinmetz et al., 1994). Whereas enhancement could correspond to a selection process, suppression could be a correlate of shifts of attention or inhibition of return. These hypotheses need to be tested experimentally.

A correlate of this focal allocation of attention can be seen in humans. A recent fMRI experiment describes a parietal activation in a task that requires subjects to attend to a specific spatial location in order to detect a subsequent visual stimulation at this same location (Sereno et al., 2001). This parietal region is located in the superior parietal cortex, just beyond the medial tip of the intraparietal sulcus and the

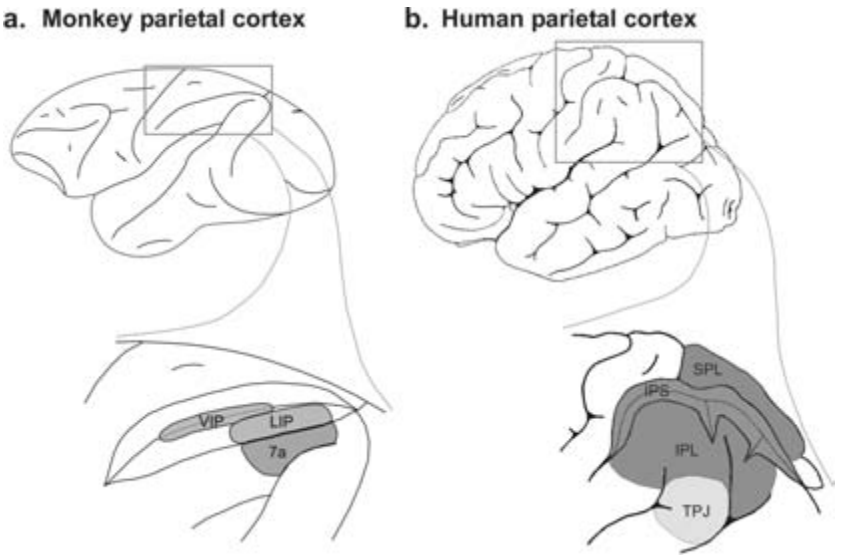


Figure 14.1
Parietal brain regions activated by attention are represented (*left*) in monkeys, area 7a, on the inferior parietal convexity, the lateral intraparietal area LIP, on the lateral bank of the intraparietal sulcus (opened up for the purpose of the figure), and the ventral intraparietal area VIP on the fundus of the intraparietal sulcus; and (*right*) in humans: the intraparietal sulcus is opened up and the superior parietal lobule (SPL), the intraparietal sulcus (IPS), the inferior parietal lobule (IPL), and the temporoparietal junctions (TPJ).

authors suggest that it could be the homologue of LIP in monkeys. Because the BOLD signal measured in fMRI studies is a metabolic marker, it cannot be inferred whether this parietal activation corresponds to an enhancement of the neuronal responses or a suppression, as both of these processes are glucose-dependent.

Baseline Attentional Modulations

Spatially selective attention also modulates the baseline spike rate of neurons. This phenomenon is mostly described as an enhancement and is thought to reflect an anticipation of the stimulus appearance and the voluntary direction of attention toward the location of behavioral significance (Colby et al., 1996). It is shown in several parietal areas as well as in other areas such as extrastriate area V4 or the frontal eye field (FEF). Because it takes place in the absence of any sensory stimulation, this baseline modulation cannot be interpreted as a gating signal for sensory inputs.

A recent fMRI study has described this expectancy signal in humans (Kastner et al., 1999). The subjects had to attend a particular location in space and to count the occurrences of a specific visual stimulus at that location. The analysis of cerebral activity evoked during expectancy periods in the absence of any visual stimulation highlighted several cortical regions. The superior parietal lobule and the intraparietal sulcus were highly activated, and less consistently the inferior parietal lobule. The frontal cortex and the occipital visual areas were also activated. The authors suggest that a frontoparietal attentional network is at the origin of a top-down signal that biases the neural activities in extrastriate areas in favor of the attended location, in the absence of any visual stimulation.

While both of these modulations reflect the neuronal processes responsible for attentional functions, they do not provide a clear picture of how these processes account for the different functional manifestations of attention. In the following, we will review the physiological evidence relative to specific aspects of selective attention mechanisms in both humans and monkeys.

Exogenous Involuntary and Endogenous Voluntary Shifts of Visual Attention

Two different types of selective attention are classically defined: an exogenous allocation of attention, involuntarily driven by the intrinsic characteristics of an external stimulus, and an endogenous allocation of attention, driven by voluntary "top-down" cognitive processes. Exogenous shifts of attention are often studied with a cued peripheral detection task also known as an exogenous Posner task (Posner, 1980). In such a task, a cue indicates the probable location of the target to be detected as quickly as possible. When the target appears at the cued location (valid condition), both humans and animals have very fast detection times (Posner, 1980; Robinson et al., 1995). When the target appears at another location than the cued location (invalid condition), detection times are slower. This cost is interpreted as the time needed by the subject to reorient his attention toward the target.

Monkey parietal neurons have been shown to contribute to these automatic shifts of attention. For example, in an exogenous Posner task, more than a third of the neurons show either an enhancement or a suppression of their response when a predictable visual stimulus falls in their receptive field (Robinson et al., 1995). Using the same paradigm, Davidson and Marrocco also describe changes in neuronal response latencies as a function of cue validity (Davidson & Marrocco, 2000). When the target appears at a validly cued position inside the receptive field, the latency of the neuronal response is shorter than when the target is invalidly cued away from the receptive field. When scopolamine, a cholinergic antagonist, is injected at the recording site of these neurons, reaction times to the detection of the target at the cued loca-

tion are significantly slowed down. All this taken together strongly suggests a role of monkey parietal cortex in the covert automatic shifts of attention.

As for voluntary shifts of attention, these are usually tested using a symbolic central cue that orients attention to a given spatial location. To our knowledge, this has never been tested directly in monkeys. A recent report that manipulates endogenous attention in monkeys is that of Gottlieb and colleagues, who use a variant of a match to sample task, in which a central cue orients the monkey's attention toward a given item (and thus position) of a stable visual array (Gottlieb et al., 1998). They show that the cue induces an attentional modulation on the response of neurons of the intraparietal cortex to the matching visual stimulus of the array that is already present in their receptive field. In the context of this task, it is difficult to interpret this modulation as a purely attentional signal as the task of the monkey is to prepare an eye movement toward this target. But because this enhancement is observed only when a target is present at the locus of attention, the authors interpret it as the selection of a location in a map encoding saliency and a signature for spatial allocation of attention. Thus monkey parietal neurons appear to be modulated both by exogenous and endogenous attentional shifts.

These two aspects of attention have been investigated both in patients with parietal lesions and in healthy persons, using neuroimaging techniques. Patients with parietal lesions are impaired in a Posner task (Posner et al., 1984), but in a somewhat different way than monkeys. In the valid condition, patients exhibit no particular difficulty in shifting their attention toward the ipsilesional or contralesional visual fields in spite of a decreased reaction time for detecting the target. By contrast, they are impaired in the invalid condition, when the cue has been presented on the ipsilesional side and the target has to be detected on the contralesional side of space. According to the authors, this suggests that the patients are not impaired in shifting their attention but rather in disengaging it from the ipsilesional space. The fact that parietal patients are more impaired on invalid cue conditions, while monkeys are more affected in the valid cue condition during scopolamine injections, implies that the role of parietal cortex in automatic attentional shifts is different between humans and monkeys.

On a centrally cued Posner task, driven by endogenous attention, parietal patients are, like in an exogenous Posner task, slower to respond to a invalidly cued target, when it appears in their contralesional visual field (Posner et al., 1984). These results are also interpreted in terms of a disengagement deficit. An fMRI comparison between the networks activated during exogenous and endogenous attentional shifts has been carried out in normal human subjects (Rosen et al., 1999). This study shows that the same network is involved in both conditions, which includes in the parietal cortex the superior parietal lobule, the inferior parietal lobule, and the temporoparietal

junction (also activated in a neutral noncued condition). Because the patients tested by Posner et al. (1984) often have very large lesions, it is difficult to correlate their observations specifically with one of the parietal regions activated in the Rosen et al. study. The presence of a mild neglect nonetheless suggests an involvement of the temporoparietal region (see, for example, Mort et al., 2003). Another recent study has carried out a precise mapping of the cortical regions activated during the valid and invalid conditions of a voluntary orienting of attention, using event-related fMRI (Corbetta et al., 2000). By selectively analyzing the activation due to cue presentation and to target presentation, the authors show an intraparietal activation during the cue presentation suggesting a role of this area in orienting attention, and an activation of the temporoparietal junction during target presentation, especially when the target falls at an unexpected location, suggesting a role of this region in exogenous attentional shifts. These observations refine those of Rosen et al. and are interpreted in terms of two parietal functional systems: the intraparietal cortex for voluntary orienting of attention and the temporoparietal junction for the automatic orienting of attention.

Using a different task, researchers have tried to further dissect the underlying cortical regions implicated in the voluntary orienting of attention as compared to voluntarily maintaining of attention at a given spatial location (Yantis et al., 2002). The authors find that both the superior and inferior parietal lobules are activated in correlation with the voluntary shift of attention from one side of space to the other, whereas a posterior ventral section of the intraparietal sulcus is specifically activated when attention has to be maintained at a given location. fMRI studies thus seem to suggest that different human parietal areas are involved in different aspects of attentional control. Further experiments will have to be carried out to address these different open questions regarding exact implication of each of the human and monkey parietal cortices in voluntary and automatic attentional shifts.

Because selective spatial attention is closely related to visual exploration, the link between attentional shifts and eye shifts has drawn a lot of interest. In particular, Rizzolatti et al. have put forward the premotor theory of attention, which stipulates that the orientation of attention uses the same network as that responsible for eye orientation, except for the very last motor steps (Rizzolatti et al., 1987). The fact that in monkey parietal cortex, the majority of saccadic neurons is modulated by attentional states and vice versa supports this suggestion (see, for example, Colby et al., 1996). Interestingly, the parietal networks activated during voluntary attentional shifts and voluntary eye movements appear to be closely overlapping in humans, including both the intraparietal cortex and the temporoparietal junction, with a notable greater activation of the right parietal cortex during attentional shifts (Corbetta et al., 1998). However, as will be seen below, recent data by Wardak and colleagues show that

reversible parietal inactivation can lead to significant deficits in covert attentional tasks in the absence of any observable saccadic deficit. Thus, further work needs to be carried out to test the validity of the premotor theory of attention in its original formulation and to characterize the relationship between the attentional and saccadic networks in both species.

Saliency and Visual Competition

The capacity of an object to draw one's attention is known as its saliency. This property is a function of both intrinsic properties (abrupt onset, color) and extrinsic properties (familiarity, behavioral significance). In the previously mentioned study, Gottlieb and colleagues very elegantly show that intraparietal LIP neurons are modulated by the saliency of visual stimuli (Gottlieb et al., 1998). The authors used stable visual arrays whose items initially did not fall onto the receptive field of the neuron being recorded. The monkey is then required to make a saccade that brings a specific item of the stable array into the receptive field. Their first interesting observation is that bringing one of the stable stimuli into the receptive field of the neurons hardly elicits any response, suggesting that these neurons do not encode in their response nonnovel, nonrelevant visual information. However, when the monkey is, in a second step, centrally cued to make a second saccade toward the stimulus that is now present into the receptive field—thus making it relevant to the monkey—the neuron's discharge is highly increased and maintained up to the execution of the relevant behavior. Thus LIP neurons respond selectively to relevant stimuli, as well as to abrupt onset stimuli, but not to nonnovel, nonrelevant stimuli, suggesting that they encode saliency rather than purely visual or motor-preparation information. This idea that LIP acts as a saliency map is further confirmed by an elegant study in which the authors show the existence in this area of a neural signal reflecting the spatial allocation of attention as measured psychophysically (Bisley & Goldberg, 2003a, 2003b). Another study in the inferior parietal lobule of the monkey, in area 7a, shows that it specifically encodes the location of a salient stimulus amongst distracters when saliency is determined by the color of the stimuli (Constantinidis & Steinmetz, 2001). These two parietal regions are thus believed to encode what is known as a saliency map, that is, a map of the sensory world in which the representation of the outstanding stimuli is enhanced with respect to the other stimuli (Gottlieb et al., 1998).

However, the notion of saliency implies a preprocessing or filtering of the visual scene among which only a few items will be selected as salient. Hence the idea that there is a competition between putative targets of attention and distracters, taking place along the visual extrastriate processing stream, that will finally lead to the emerging of a saliency map in the higher level areas (Desimone & Duncan, 1995). In this

context, the parietal cortex has been suggested to act as a controller, sending a top-down biasing signal to favor an object or a location on the low level sensory maps, and/or to be involved in accessing the saliency map and filtering the distractors according to the behavioral requirements. Evidence of this is found in the description of a bilateral parietal patient who exhibits difficulties in processing visual targets when these are surrounded by distractors, all the more when their saliency is high (Friedman-Hill et al., 2003). This patient has an intact representation of the targets but finds it difficult to filter them from the surrounding distractors. Although these deficits are very similar to those observed in monkeys with lesions in extrastriate area V4, where preattentive competitive processes are thought to take place, their interpretation in terms of filtering deficits fits best the presumed role of parietal cortex as a controller.

Another level of cognitive processing at which competition in the selection of visual targets by attention also expresses itself is interhemispheric competition. The best illustration of this phenomenon is the fact that some parietal human patients with mild or no neglect symptoms are unable to perceive consciously the contralateral target of two simultaneously flashed stimuli, one in each hemifield. This deficit is called extinction and can be overdriven by giving a temporal advantage of about 200 msecs to the contralateral target (Rorden et al., 1997).

Monkeys with extended lesions including the cortex from the intraparietal sulcus to the superior temporal and the lunate sulci (thus, both areas LIP and 7a) show very similar deficits both in the visual and in the somatosensory modalities (Deuel & Farrar, 1993). In our laboratory, we have specifically tested the involvement of area LIP in extinction (Wardak et al., 2002). We reversibly inactivated the electrophysiologically defined area LIP with muscimol, a GABAa agonist. Monkeys were presented with either one or two competing visual stimuli, one in each hemifield, and had to make a saccade towards one of them. On dual presentations, a temporal delay was introduced between the contralesional target and the ipsilesional one, ranging from –40 msecs (ipsilesional leading) to +280 msecs (contralesional leading) with a step of 40 msecs. Following unilateral muscimol injections, monkeys exhibit a biased choice toward the ipsilesional target, which is cancelled only when the contralesional target leads the ipsilesional one by more than 180 msecs on average, which is very close to what is seen in human patients (figure 14.2a). Interestingly, when only one target was presented (50 percent of all the trials), the monkey missed a high percentage of those presented contralaterally, whereas no specific deficit was ever observed on a regular saccade task toward contralateral targets (figure 14.2b). This last observation highly suggests that the deficits observed in the extinction task reflect lesion-induced asymmetries in interhemispheric competition processes. This type of competition undoubtedly contributes to the construction of saliency maps. However, it is most likely very different in nature from the preattentive bottom-up competitive processes described in lower visual areas

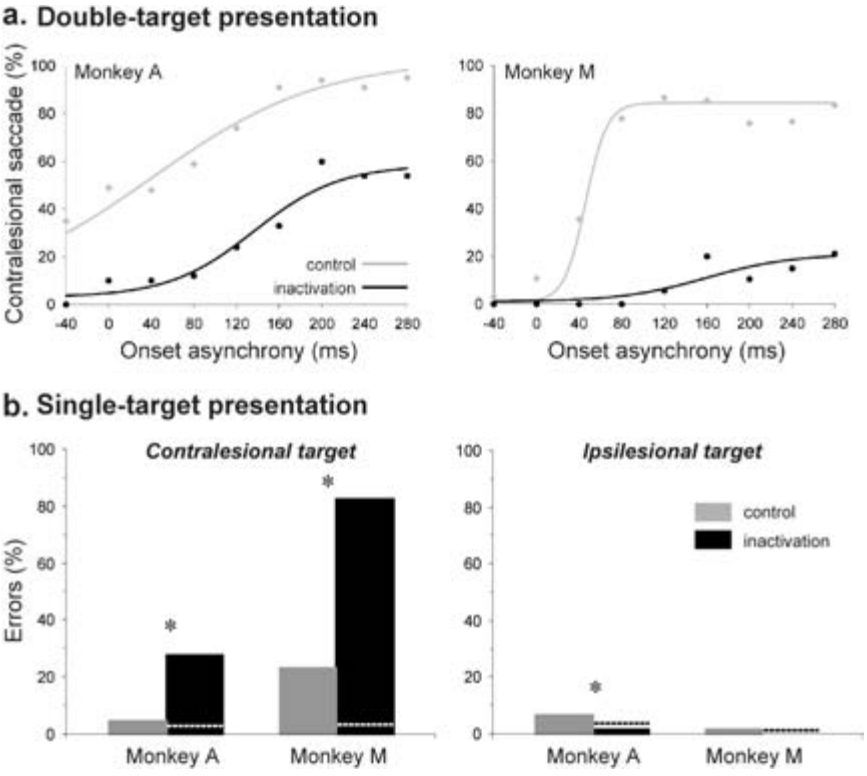


Figure 14.2

Saccades to bilateral and unilateral targets during reversible LIP inactivation, in an extinction task. (*a*) Proportion of contraversive saccades on double-target presentations in monkeys M and A for each target onset asynchrony, and corresponding logistic regression fits through inactivation and control data. Positive asynchrony means contralesional target leading the ipsilesional one. (*b*) Omissions on single-target trials intermingled with double-target presentations. Horizontal dashed lines indicate the percentage of single-target trials on which the monkeys failed to respond in a standard visually guided saccade task. Gray bars correspond to control data, black bars correspond to data after LIP inactivation. *$p < 0.05$.

such as V4, and which take place at the level of the single cell receptive fields. To our knowledge, this interhemispheric competition has not yet been addressed as such in human studies.

Visual Search

Visual search is a very popular paradigm in the study of attention. However, it is very complex in that such different aspects as the displacement of attention, its time course, its triggers, and so on cannot be easily dissociated. In its most standard versions, the subject has to find a predefined target in an array of distractors which are different from the target by at least one of their visual features (*feature search*, for example, of a green triangle in the midst of yellow triangles), or else, which are a combination of several visual features of the target (*conjunction search*, for example, of a green triangle in the midst of green squares, yellow triangles, and yellow squares). It has been found that in a feature search, the time needed to find the target is not related to the number of distractors, while in the conjunction search there exists a linear relationship between reaction times and the number of distractors. The pioneering studies in this field interpreted this distinction as evidence for a parallel preattentive processing of the target in the feature search, and a serial processing of space by attention to achieve the binding of the features of the different items present in the visual field in the conjunction search (Treisman & Gelade, 1980).

In this theoretical context, the parietal cortex has been suggested to play an important role in this putative process of feature binding. In support of this, fMRI studies have shown a superior parietal lobule activations in visual conjunction but not feature search (Corbetta et al., 1995). Also, unilateral parietal patients with neglect exhibit deficits in reaction times specifically for conjunction search (Eglin et al., 1989). Biparietal patients with Balint's syndrome show deficits in combining the different attributes of objects and exhibit a particularly high rate of illusory conjunctions (Friedman-Hill et al., 1995). TMS stimulations on the parietal cortex also lead to specific impairment on conjunction visual search (Ashbridge et al., 1997).

However, since the first studies by Treisman and colleagues, it has been shown that if the feature task is made more difficult, for example by reducing the difference between the target and the distractors or by having nonhomogeneous distractors, reaction times start to increase with the size of the distractor set. In view of these observations, a continuum in search efficiency has been proposed, requiring variable combinations of parallel and serial processing depending on search difficulty (Duncan & Humphreys, 1989). As all the above studies used very easy feature search, task difficulty effects cannot be ruled out.

Recent fMRI studies have tested again the implication of parietal cortex in visual search using both easy and difficult conjunction and feature search conditions. They

show that intraparietal sulcus and superior parietal lobule activation is correlated with search difficulty rather than with feature binding per se (Nobre et al., 2003; Donner et al., 2002). Both these studies also identify a small binding specific region but not at the same location: Donner et al. locate it in the temporoparietal junction, whereas Nobre et al. locate it in the superior parietal lobule.

This correlation between parietal activation and search difficulty does not say anything on whether attentional shifts during visual search are serial or parallel. Nobre et al. show a strong correlation between parietal activation and reaction times suggesting serial attentional shifts, while a previous study had not found any such effects (Leonards et al., 2000). The contribution of parietal cortex to any of these two different processing modes thus remains an open question.

In monkeys, very few studies if any have investigated the role of parietal cortex in visual search. We have recently studied the involvement of area LIP in this task using reversible lesions, testing overt visual search, allowing both attentional and eye shifts (Wardak et al., 2002), and covert visual search, allowing attentional shifts exclusively. We found that LIP inactivations induce an increase in the time and in the number of saccades necessary to find the target in a conjunction visual search when the latter is in the contralateral hemifield (figure 14.3a). The increase in search time is due in large part to an ipsilateral bias in the exploratory strategy when the stimuli are distributed evenly between the ipsi- and contralesional visual hemispaces (figure 14.3b). However, because this deficit persists when the search array is located in the contralateral hemifield, it cannot be accounted for by a simple motor bias or by interhemispheric attentional competition factors (figure 14.4). Instead, the deficit seem to imply an impaired ability to select a target from among neighboring competitors, reminiscent of the above-mentioned observations made by Friedman-Hill et al. (2003) in a biparietal patient. Inactivations did not induce any deficits on easy feature search tasks. Because we did not test difficult feature search conditions, we cannot conclude to a specific role of LIP in binding or in difficult attentional tasks.

In a covert version of the visual search task, in which the monkey had to respond by a lever press whenever he identified the target in rapidly and successively presented search arrays, an increased reaction time was observed for detecting a contralesional target in conjunction search but also in feature search (Wardak et al., 2004; figure 14.5). The magnitude of the deficits were about the same in conjunction and difficult feature search conditions, but were considerably reduced in easy feature search conditions. Deficits were thus dependent on the saliency of the visual target amongst the distractors. Thus, area LIP does not seem to contribute crucially to feature binding.

We also tested the contribution of LIP to either of parallel or serial processing. A signature of a role of LIP in the latter should be a significant correlation between the amplitude of deficit and the number of distractors. No such correlation was found.

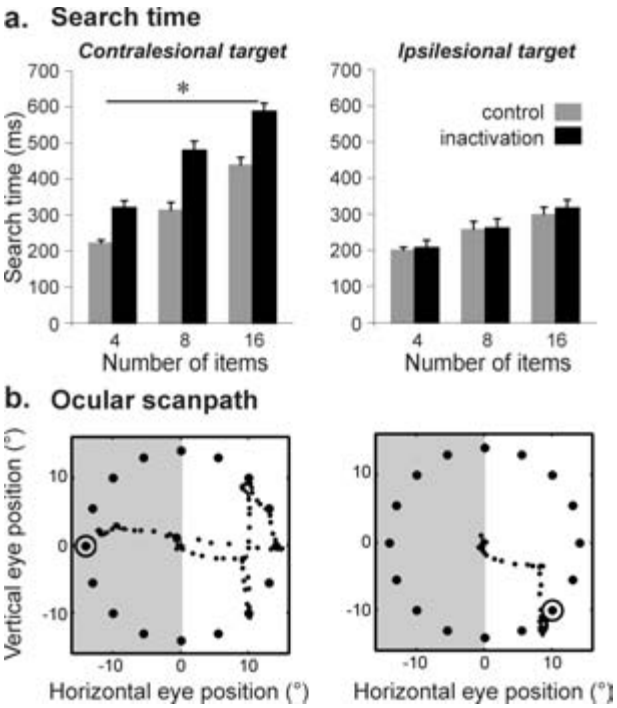


Figure 14.3

Visual search performance during reversible LIP inactivation. (*a*) Search time is shown as a function of display size for contralesional (*left*) and ipsilesional targets (*right*). Search time for a target defined by a conjunction of color (red versus orange) and form (circle versus square) is defined as the interval between the search pattern onset and the start of the saccade landing on the target (bar height is mean search time, error bar is standard error). Gray bars correspond to control data, black bars correspond to data after LIP inactivation. *corrected $p < 0.05$. (*b*) Single trial examples of visual search patterns from Monkey M during LIP inactivation. The smallest dots represent eye position sampled every 4 msecs, large dots the search stimuli and the circle the location of the target. The contralesional field is on the left.

Although this questions the role of LIP in serial shifts of attention, the fact that monkeys are often overtrained, makes it difficult to assess the exact contributions of serial and parallel processing to the task, even in the conjunction condition. Further manipulations of the search condition are necessary to understand the role of LIP in covert attentional shifts and confirm its functional homology with the human intraparietal region specifically activated by attentional processes. Interestingly, these deficits in tasks requiring covert attention are observed in the absence of saccadic deficits. This calls to a reevaluation of the premotor theory in terms of the cortical network that contributes to it.

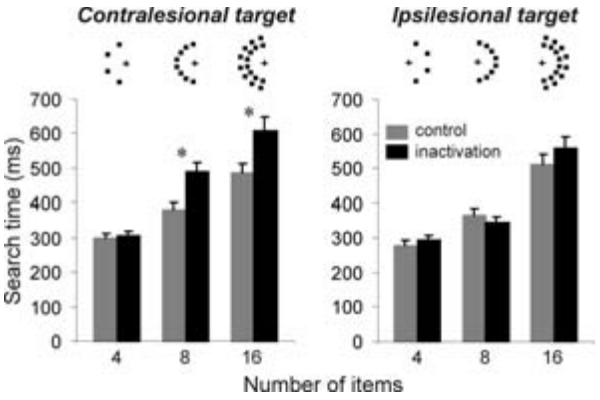


Figure 14.4
Visual search performance for monkey M during reversible LIP inactivation, when all items are placed within the same hemifield. Search time for a target defined by a conjunction of color (yellow versus green) and form (diamond versus square) is defined as the interval between the search pattern onset and the start of the saccade landing on the target (bar height is mean search time, error bar is standard error). Spatial layout of items is represented above each display size for the contralesional on the left and ipsilesional side on the right. The contralesional field is on the left of each display. Gray bars correspond to control data, black bars correspond to data after LIP inactivation. *corrected $p < 0.05$.

Conclusion

In conclusion, it appears that despite the complexity of attentional processes, their manifestations seem at first sight similar for similar tasks in both humans and monkeys and repeatedly activate well-defined elements of a parietal attentional network. In humans, these regions are essentially the superior parietal lobule, the intraparietal sulcus, and the parietotemporal junction. In monkeys, the attentional regions identified to this date by electrophysiology are LIP and 7a, but the forthcoming effort of using fMRI techniques on the study of monkey attentional networks might reveal other parietal areas of interest. We have shown that LIP inactivations produce some of the deficits observed in neglect syndrome in humans. Not all symptoms of neglect were found however. Recent studies suggest that human homologue of LIP, as defined by its saccadic and attentional responses is located in the intraparietal sulcus, in a somewhat similar position as in the monkey (Sereno et al., 2001). However, as seen above, whereas it seems that exogenous and endogenous attentional shifts activate identical areas in monkeys (LIP and 7a), a segregation of these two attentional processes has been described in humans, exogenous attention activating mostly the parietotemporal junction and endogenous attention activating the intraparietal sulcus. In the light of current knowledge, and in spite of the apparent similarities, it

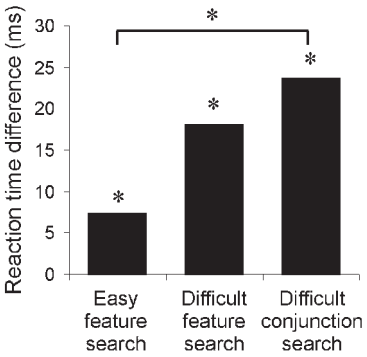


Figure 14.5

Covert visual search performance of monkey M during reversible LIP inactivation. Results are expressed in terms of reaction time difference between the control and the inactivation conditions. Easy and difficult search conditions as well as conjunction search condition are represented. *corrected $p < 0.05$.

is difficult to suggest strict homologies between attention-activated parietal regions in humans and monkeys. Indeed, in humans, the parietotemporal junction is part of this attentional network in addition to the intraparietal sulcus. Such an area is still not identified in monkeys. Area 7a could be a candidate for this region, although very few studies have probed its implication in attentional shifts. fMRI in monkeys will be a key technique to answer this question. Second, the segregation between areas involved in automatic and voluntary attentional shifts described in humans does not seem to exist in monkeys in the light of the available data. Again, fMRI in monkeys, followed by well-designed electrophysiological and inactivation experiments, will enable to answer these questions. Such studies will be crucial in continuing the effort of probing the similarities between the human and monkey brain, and in the context of attention, possibly probing evolutionary divergences in the organization of a cognitive cortical network.

Here we have considered only attention-activated parietal regions. However, the attentional network involves a more complex parieto-fronto-subcortical network in both humans and monkeys. More experimental work needs to be carried out in monkeys to uncover the extent of this network and the functional interactions within it.

Acknowledgments

The results presented here were made possible thanks to the financial contribution of the Fondation pour la Recherche Médicale, the Fyssen Foundation, and the French Ministry of Research.

References

Ashbridge, E., Walsh, V., et al. (1997). Temporal aspects of visual search studied by transcranial magnetic stimulation. *Neuropsychologia*, *35*(8), 1121–1131.

Ben Hamed, S., Duhamel, J. R., et al. (2002). Visual receptive field modulation in the lateral intraparietal area during attentive fixation and free gaze. *Cerebral Cortex*, *12*(3), 234–245.

Bisley, J. W., & Goldberg, M. E. (2003a). Neuronal activity in the lateral intraparietal area and spatial attention. *Science*, *299*(5603), 81–86.

Bisley, J. W., & Goldberg, M. E. (2003b). The role of the parietal cortex in the neural processing of saccadic eye movements. *Advances in Neurology*, *93*, 141–157.

Bushnell, M. C., Goldberg, M. E., et al. (1981). Behavioral enhancement of visual responses in monkey cerebral cortex. I. Modulation in posterior parietal cortex related to selective visual attention. *Journal of Neurophysiology*, *46*(4), 755–772.

Colby, C. L., Duhamel, J. R., et al. (1996). Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. *Journal of Neurophysiology*, *76*(5), 2841–2852.

Constantinidis, C., & Steinmetz, M. A. (2001). Neuronal responses in area 7a to multiple-stimulus displays: I. neurons encode the location of the salient stimulus. *Cerebral Cortex*, *11*(7), 581–591.

Corbetta, M., Akbudak, E., et al. (1998). A common network of functional areas for attention and eye movements. *Neuron*, *21*(4), 761–773.

Corbetta, M., Kincade, J. M., et al. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience*, *3*(3), 292–297.

Corbetta, M., Shulman, G. L., et al. (1995). Superior parietal cortex activation during spatial attention shifts and visual feature conjunction. *Science*, *270*(5237), 802–805.

Davidson, M. C., & Marrocco, R. T. (2000). Local infusion of scopolamine into intraparietal cortex slows covert orienting in rhesus monkeys. *Journal of Neurophysiology*, *83*(3), 1536–1549.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Reviews of Neuroscience*, *18*: 193–222.

Deuel, R. K., & Farrar, C. A. (1993). Stimulus cancellation by macaques with unilateral frontal or parietal lesions. *Neuropsychologia*, *31*(1), 29–38.

Donner, T. H., Kettermann, A., et al. (2002). Visual feature and conjunction searches of equal difficulty engage only partially overlapping frontoparietal networks. *NeuroImage*, *15*(1), 16–25.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458.

Eglin, M., Robertson, L. C., et al. (1989). "Visual search performance in the neglect syndrome." *Journal of Cognitive Neuroscience*, *4*: 372–381.

Friedman-Hill, S. R., Robertson, L. C., et al. (2003). Posterior parietal cortex and the filtering of distractors. *Proceedings of the National Academy of Sciences*, *100*(7), 4263–4268.

Friedman-Hill, S. R., Robertson, L. C., et al. (1995). Parietal contributions to visual feature binding: evidence from a patient with bilateral lesions. *Science*, *269*(5225), 853–855.

Gottlieb, J. P., Kusunoki, M., et al. (1998). The representation of visual salience in monkey parietal cortex. *Nature*, *391*(6666), 481–484.

Kastner, S., Pinsk, M. A., et al. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, *22*(4), 751–761.

Leonards, U., Sunaert, S., et al. (2000). Attention mechanisms in visual search: An fMRI study. *Journal of Cognitive Neuroscience*, *12 Suppl 2*, 61–75.

Mort, D. J., Malhotra, P., et al. (2003). The anatomy of visual neglect. *Brain*, *126*(Pt 9), 1986–1997.

Mountcastle, V. B., Andersen, R. A., et al. (1981). The influence of attentive fixation upon the excitability of the light-sensitive neurons of the posterior parietal cortex. *Journal of Neuroscience*, *1*(11), 1218–1225.

Nobre, A. C., Coull, J. T., et al. (2003). Brain activations during visual search: Contributions of search efficiency versus feature binding. *NeuroImage*, *18*(1), 91–103.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25.

Posner, M. I., Walker, J. A., et al. (1984). Effects of parietal injury on covert orienting of attention. *Journal of Neuroscience*, *4*(7), 1863–1874.

Rizzolatti, G., Riggio, L., et al. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia*, *25*(1A), 31–40.

Robinson, D. L., Bowman, E. M., et al. (1995). Covert orienting of attention in macaques. II. Contributions of parietal cortex. *Journal of Neurophysiology*, *74*(2), 698–712.

Rorden, C., Mattingley, J. B., et al. (1997). Visual extinction and prior entry: Impaired perception of temporal order with intact motion perception after unilateral parietal damage. *Neuropsychologia*, *35*(4), 421–433.

Rosen, A. C., Rao, S. M., et al. (1999). Neural basis of endogenous and exogenous spatial orienting: A functional MRI study. *Journal of Cognitive Neuroscience*, *11*(2), 135–152.

Sereno, M. I., Pitzalis, S., et al. (2001). Mapping of contralateral space in retinotopic coordinates by a parietal cortical area in humans. *Science*, *294*(5545), 1350–1354.

Steinmetz, M. A., Connor, C. E., et al. (1994). Covert attention suppresses neuronal responses in area 7a of the posterior parietal cortex. *Journal of Neurophysiology*, *72*(2), 1020–1023.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive, Psychology*, *12*(1), 97–136.

Wardak, C., Olivier, E., et al. (2002). Saccadic target selection deficits after lateral intraparietal area inactivation in monkeys. *Journal of Neuroscience*, *22*(22), 9877–9884.

Wardak, C., Olivier, E., Duhamel, J.-R. (2004). A deficit in covert attention after parietal cortex inactivation in the monkey. *Neuron*, *42*, 501–508.

Yantis, S., Schwarzbach, J., et al. (2002). Transient neural activity in human parietal cortex during spatial attention shifts. *Nature Neuroscience*, *5*(10), 995–1002.

IV Cognitive Control and the Frontal and Cingulate Cortices

15 The Rostral-Caudal Axis of Cognitive Control within the Lateral Frontal Cortex

Michael Petrides

Patients with cortical excisions of the lateral frontal cortex perform well on many standard tests of memory and on various other cognitive tests (Petrides, 2000a; Stuss et al., 1986). Damage to the lateral frontal cortex, however, results in impairments on several memory and other cognitive tasks when adequate performance depends on the expression of various executive processes (Petrides, 2000a; Shallice & Burgess, 1996). It is now well established that the frontal cortex plays an important role in the top-down regulation of cognitive and behavioral processes (Luria, 1969; Petrides, 1996, 2000a; Robbins, 1996; Shallice & Burgess, 1996). The frontal cortex, however, is not a homogeneous region of the cerebral mantle but comprises several distinct areas that differ in terms of their architecture (Brodmann, 1905, 1908, 1909; Economo & Koskinas, 1925; Walker, 1940; Barbas & Pandya, 1987, 1989; Petrides & Pandya, 1994). Furthermore, these distinct architectonic areas have been shown to have different afferent and efferent connections with other parts of the brain (see Petrides & Pandya, 2002). These anatomical facts suggest that the different architectonic areas are likely to be involved in distinct aspects of the higher-level control of cognitive processing and behavior that is the domain of frontal function.

In trying to understand the functional organization of the lateral frontal cortex, we have studied the cognitive effects of lesions of the frontal cortex in both human patients and macaque monkeys in a comparative manner. There were two main reasons for this strictly comparative approach. First, the aim was to uncover the fundamental principle of functional organization of higher order control processing in the primate frontal cortex. Since the fundamental organizational scheme is likely to be the same across all primate brains, a comparative approach is more likely to reveal the essential aspects of frontal cortex organization. Second, since lesions in patients are rarely restricted precisely to particular anatomically defined architectonic areas of the frontal cortex, while in the monkey such lesions can be made with great precision, research on monkeys can establish dissociations in the functional contributions of various sectors of the frontal cortex that can only imperfectly be studied in the human brain. This comparative work has revealed a dorsal-ventral axis in the

organization of executive control in the mid-lateral parts of the prefrontal cortex and also a rostral-caudal axis of executive control. The dorsal versus ventral distinction in executive control, which was proposed in the early 1990s (Petrides, 1994, 1996), has been the subject of recent reviews (Petrides, 2000b, 2000c) and therefore the present chapter will focus on the rostral-caudal axis in the organization of executive control within the lateral frontal cortex. This distinction was established in the 1980s from monkey lesion studies that examined differences in the effects of mid-dorsolateral frontal lesions and caudal dorsolateral frontal lesions in conditional learning and working memory (see Petrides, 1987, for an early review).

The most caudal part of the frontal lobe comprises the classic motor region that is organized in somatotopic maps (e.g., Penfield & Boldrey, 1937; Grafton, Hari, & Selenius, 2000). This classic motor region lies on the precentral gyrus in the human brain and includes the motor cortical area 4 and the caudal premotor area 6 (figure 15.1). In the human brain, area 4 is found mostly in the rostral bank of the central sulcus and at a variable part of the caudal precentral gyrus (e.g., Economo & Koskinas, 1925; Geyer et al., 1996). The remaining rostral part of the precentral gyrus is occupied by caudal area 6. Immediately anterior to the motor region lies the caudal dorsolateral frontal region that occupies the caudal part of the superior and middle frontal gyri in the human brain. This caudal dorsolateral frontal region comprises the rostral part of area 6 and area 8. Further rostrally, lies the mid-dorsolateral prefrontal region (areas 9, 46 and 9/46) that occupies the midsector of the superior and middle frontal gyri (figure 15.1).

Recent investigations showed that the classical motor region is formed by a mosaic of areas. Some of these areas (e.g., area 4) are involved in fine motor control, while others (e.g., caudal area 6) in sensorimotor transformations for reaching, grasping, and manipulation of objects (He, Dum, & Strick, 1993; Grafton et al., 2000; Picard & Strick, 2001; Rizzolatti, Luppino, & Matelli, 1998; Rizzolatti & Luppino, 2001). The functional organization of this part of the frontal lobe will not concern us further in this article. The present chapter focuses on the differences between the mid-dorsolateral and the caudal dorsolateral frontal cortex.

Caudal Dorsolateral Frontal Cortex and Conditional Operations Versus Mid-Dorsolateral Frontal Cortex and Monitoring in Working Memory

In the 1980s we carried out a series of studies that tried to shed light on the essential nature of the contribution to cognitive processing of the dorsolateral frontal cortex. These studies involved the development of novel tasks to examine the cognitive deficits in patients who had excisions of lateral frontal cortex for the treatment of pharmacologically intractable epilepsy or low-grade, small tumors. Although there were strong indications from the patient work that there might be dissociations

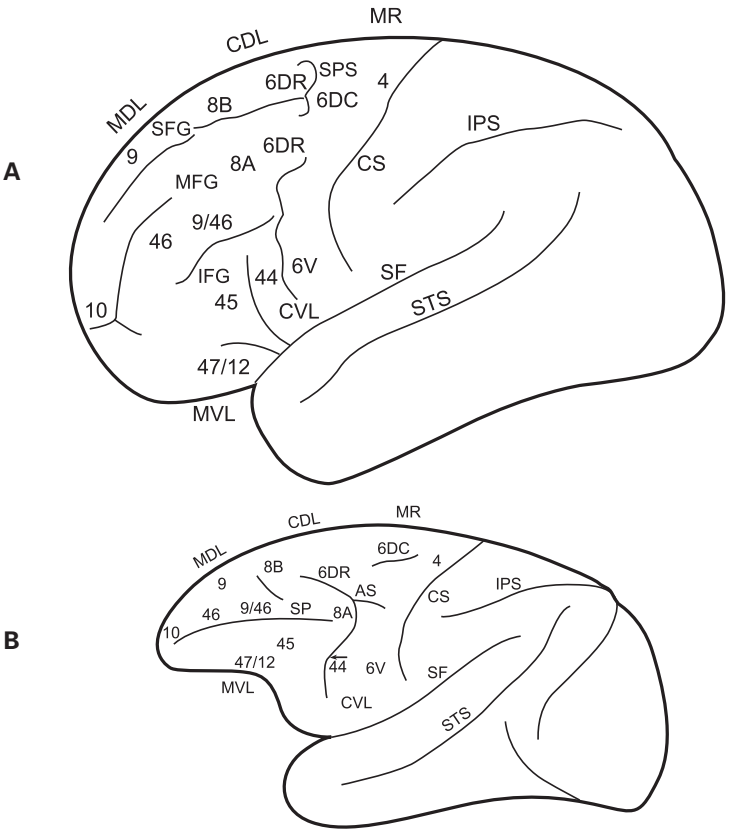


Figure 15.1
Schematic outline of the lateral surface of the human (*A*) and the monkey (*B*) left cerebral hemisphere to illustrate the major subdivisions of the lateral frontal cortex. The numbers refer to the location of the architectonic areas. The arrow above area 44 in the monkey indicates that this area is found primarily in the posterior bank of the lower branch of the arcuate sulcus. Abbreviations: AS, arcuate sulcus; CDL, caudal dorsolateral frontal cortex; CS, central sulcus; CVL, caudal ventrolateral frontal cortex; IFG, inferior frontal gyrus; IPS, intraparietal sulcus; MDL, middorsolateral frontal cortex; MFG, middle frontal gyrus; MR, motor region; MVL, midventrolateral frontal cortex; SF, Sylvian fissure; SFG, superior frontal gyrus; SP, sulcus principalis; SPS, superior precentral sulcus; STS, superior temporal sulcus.

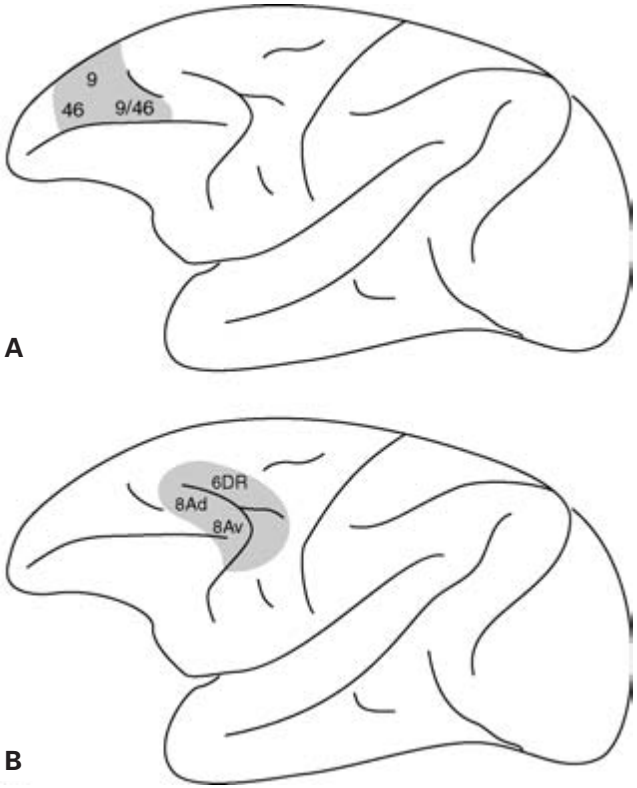


Figure 15.2
(*A*) The mid-dorsolateral (MDL) prefrontal lesion. (*B*) The caudal dorsolateral prefrontal lesion involving the dorsal periarcuate (PA) region. These lesions in the monkey were used to study the fundamental differences in the rostral-caudal axis of lateral frontal cortex organization. The numbers refer to the architectonic areas.

between the cognitive deficits that follow more rostral as compared with more caudal lesions of the dorsolateral frontal cortex, it was very difficult to establish such dissociations because there was significant overlap in the lesions. We therefore initiated a comparable series of studies on macaque monkeys. We adapted, for the monkey, the same cognitive tasks that we used with patients and examined the effects of excisions restricted to the more rostral part of the dorsolateral frontal cortex versus its more caudal part. The two lesions that we studied are shown schematically in figure 15.2. The rostral lesions involved the mid-dorsolateral frontal cortex and included cortex in the sulcus principalis and above it. Thus, these lesions involved areas 46, 9/46 and 9. The caudal dorsolateral lesions involved the cortex within the dorsal part of the

arcuate sulcus and the immediately surrounding region, namely the rostral dorsal area 6 and area 8A. We called these lesions the periarcuate lesions since they involved cortex within and surrounding the arcuate sulcus.

The studies on monkeys demonstrated striking dissociations between the cognitive sequelae of lesions to the mid-dorsolateral versus those of the caudal dorsolateral frontal cortex. Whereas lesions of the mid-dorsolateral frontal cortex caused a severe deficit on tasks that required the monitoring of information in working memory, the caudal dorsolateral lesions did not affect performance on such tasks, but yielded a massive impairment on tasks that required the selection between alternative competing responses based on conditional operations. Thus, this work on the monkey established double dissociations between the effects of mid-dorsolateral versus caudal dorsolateral frontal lesions, providing the strongest possible evidence for specialized contributions along the rostral-caudal axis of the dorsolateral frontal cortex. Some examples of this work will now be presented to illustrate these dissociations.

Our studies with patients demonstrated that dorsolateral frontal lesions yield a severe impairment on a working memory task, the self-ordered task, in which subjects are required to monitor their recent selections from a set of stimuli (Petrides & Milner, 1982). In the self-ordered task, the subjects are presented with different arrangements of the same set of stimuli and have to select a different stimulus, on each trial, until all the stimuli are selected. For instance, a stack of cards is presented to the subjects, and all cards have the same stimuli on them, except that the arrangement of these stimuli varies from card to card (figure 15.3A). The subjects have to select one of the stimuli on the top card, point to it and then turn over to the next card in order to select a different design. Thus, from the moment they start responding, the subjects must constantly compare the responses that they have already made with those still remaining to be carried out. In other words, each selection must be marked in the subject's mind and simultaneously considered in relation to the others that still remain to be selected. This process of marking mentally the stimuli that have been selected and those that remain to be selected, which we have termed "monitoring" of events within working memory, has been shown to be a critical aspect of the self-ordered task for the deficit after dorsolateral frontal lesions. For instance, the deficit after dorsolateral frontal lesions on the self-ordered monitoring task emerged against a background of normal performance on many other memory tasks that required recognition of familiar from unfamiliar stimuli and tasks that required recall from both short and long term memory (Petrides & Milner, 1982; Petrides, 2000a).

Although our studies with patients suggested that lesions that tended to involve the more rostral part of the middle frontal gyrus generated a severe impairment on this task, it was the monkey work that clearly linked the mid-dorsolateral prefrontal cortex to the monitoring operation in working memory (Petrides, 1991, 1995, 2000d). In an early study, the monkeys were trained on an analogue of the self-ordered task that

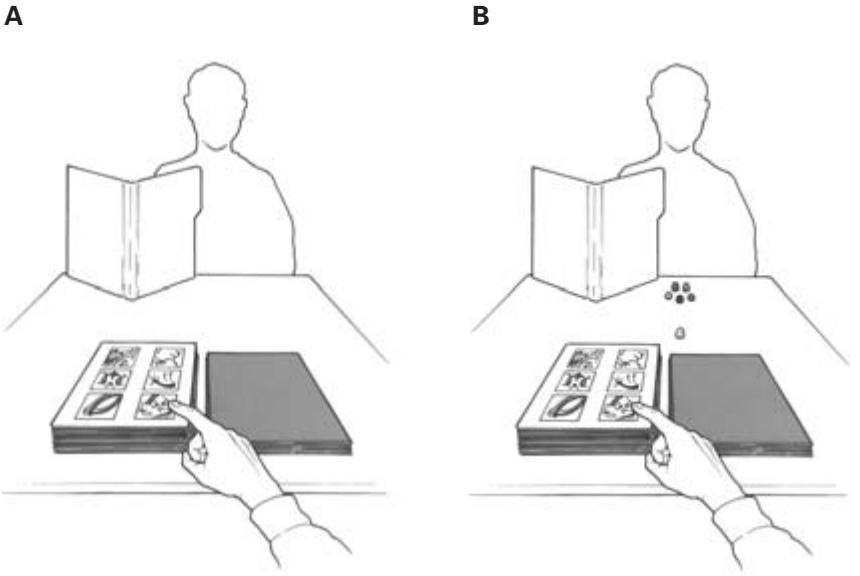


Figure 15.3
Schematic diagram of the experimental arrangement in the self-ordered monitoring working memory task (*A*) and the nonspatial conditional associative learning task (*B*) administered to patients. Note that in the conditional task (*B*) colored stimuli placed in front of the experimenter are cueing the responses. The subject has to point to the design that is the correct one for the particular color stimulus that acts as the cue in a given trial.

had been used with patients. On each trial, the monkey saw three distinct containers that differed in shape and color and also occupied a random position from trial to trial (figure 15.4). On the first trial, all three containers had a food reward hidden inside them. The animal was allowed to select one of these containers and open it to retrieve the reward. On subsequent trials, the animal was again allowed to select one of the containers and open it to find the reward. However, the reward would now be found only if the monkey opened a container that had not been selected on a previous trial. Thus, during performance of this task, the monkey must monitor within working memory his choices in order not to return to a container that has already been selected. As can been seen in figure 15.4, monkeys with mid-dorsolateral frontal lesions that had learned the task preoperatively were severely impaired when performing this task postoperatively.

A number of points need to be highlighted here. First, the animals with periarcuate lesions performed as well as the normal control animals on this task, demonstrating the specificity of the impairment along the rostral-caudal axis of dorsolateral frontal cortex. Second, the position of the three containers was randomly changed from trial

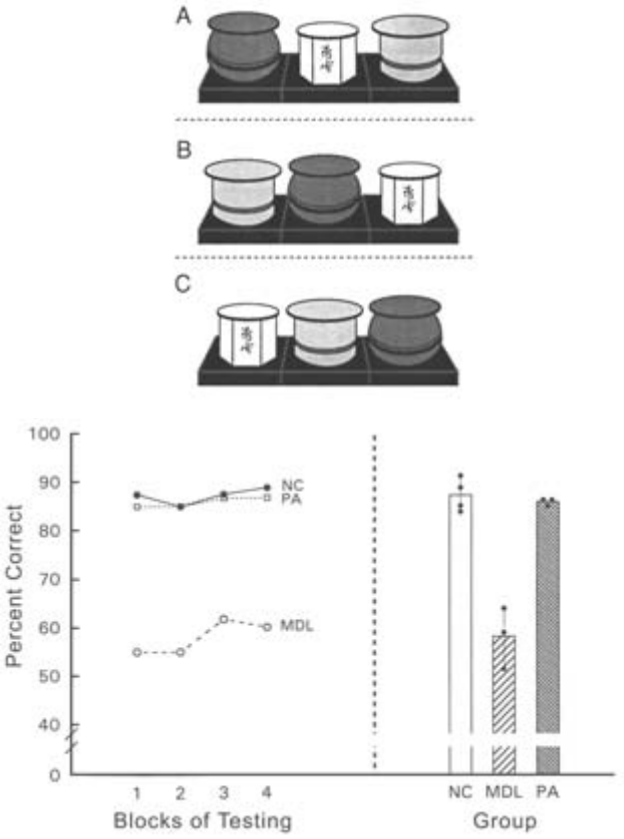


Figure 15.4

(*Upper*) Schematic diagram of the experimental arrangement in the self-ordered monitoring working memory task administered to monkeys. (*A*, *B*, and *C*) Examples of three different trials. (*Lower*) Postoperative performance of animals with mid-dorsolateral frontal lesions (MDL), animals with periarcuate lesions (PA), and normal control animals (NC). The mean percent correct performance during each block of testing (10 days per block) is shown on the left side of the lower panel. The mean percent correct performance over all four testing blocks (i.e., 40 days) is shown on the right side of the lower panel. Solid circles indicate the scores of individual animals in each group. Note that the animals with MDL lesions are severely impaired, whereas the animals with PA lesions perform as well as the NC animals.

to trial so that the animals could not have coded the stimuli in terms of their location. In order to perform well, the animals were required to select between the three different containers regardless of their location. The fact that animals with mid-dorsolateral prefrontal lesions failed this working memory task that did not require memory for spatial locations does not support the notion that the dorsolateral prefrontal cortex is selectively involved in spatial working memory (Goldman-Rakic, 1996). Third, the impairment on the nonspatial self-ordered working memory task after mid-dorsolateral prefrontal lesions was not the result of a general impairment in short-term memory because these same animals could be shown to perform normally on other memory tasks in which performance does not depend on the careful monitoring of earlier selections. For example, animals with lesions of the mid-dorsolateral prefrontal region (Petrides, 1991, 1995) or even of the entire dorsolateral prefrontal region (Bachevalier & Mishkin, 1986), just as patients with dorsolateral frontal lesions, perform normally on recognition memory tasks in which performance depends on discriminating between stimuli that were seen before and stimuli not previously seen. In addition, animals with mid-dorsolateral prefrontal lesions perform as well as normal control animals on visual short-term memory tasks in which memory is challenged by increasing the delay (i.e., storage) rather than the monitoring requirements (Petrides, 2000d).

In conclusion, this line of research demonstrated that lesions limited to the mid-dorsolateral frontal cortex (i.e., dorsal area 46 and area 9) (see figure 15.2) impair performance on working memory tasks that require monitoring of the occurrence of stimuli from an expected set (Petrides, 1991, 1995, 2000d). The analysis of the impairment of monkeys with mid-dorsolateral frontal lesions on the self-ordered and other related tasks (e.g., the externally ordered working memory tasks) is consistent with the results obtained in studies with patients. Both patients and monkeys with dorsolateral frontal lesions can remember stimuli as demonstrated by normal performance on recognition memory tests and on other short-term memory tasks. The fundamental problem on the self-ordered and other related working memory tasks stems from the *monitoring* requirements of these tasks, that is, the number of stimuli that must be kept in mind and considered as the responses are being made. The mid-dorsolateral part of the frontal cortex, therefore, appears to be a specialized area of the cerebral cortex in which information can be held on-line for monitoring (in the sense described above) and for the manipulation of stimuli. Note that the manipulation of stimuli implies the simultaneous consideration of several stimuli and thus it partly depends on monitoring. A detailed description of a model outlining this concept of the role of the mid-lateral frontal cortex in mnemonic processing has been published elsewhere (Petrides, 1994, 1996).

In sharp contrast to their normal performance on the monitoring working memory tasks that challenge animals with mid-dorsolateral frontal lesions, the monkeys with

caudal dorsolateral frontal lesions (i.e., the periarcuate lesions; see figure 15.2) perform very poorly on conditional associative tasks (Petrides, 1982, 1985a, 1986; Halsband and Passingham, 1982). These tasks require that the animal select between different possible responses in a given situation according to conditional rules, such as if A, select X, but if B, select Y. It is important to point out that patients with lateral frontal lesions invading the caudal dorsolateral frontal region also perform very poorly on such tasks. The deficit on conditional tasks in patients has been demonstrated with several types of alternative competing responses, such as hand postures (Petrides, 1985b, 1997), locations (Petrides, 1985b), and abstract designs (Petrides, 1990). For instance, in the conditional task requiring selection from a set of abstract designs (Petrides, 1990), the patients were faced with a stack of cards showing six abstract designs and six different colored stimuli (figure 15.3B). Each one of the six abstract designs corresponded, in an arbitrary manner, to one of the six colored stimuli and the patient had to learn these associations by trial and error. On each trial, the experimenter placed one of the colored stimuli in front of the others and the patient had to point to the design that was associated with that particular colored stimulus (figure 15.3B). Feedback (correct/wrong) was provided by the experimenter after each response. When an error was made, the subject pointed to other designs until she/he discovered the correct design. The next trial was initiated by placing another colored stimulus in front of the others and turning over to a new card that presented the same designs but in a new arrangement. The subject responded by pointing to the design that she/he thought was correct for this new colored stimulus. Testing continued in this manner until the subject reached a preset number of correct responses. Patients with lateral frontal lesions invading the caudal dorsolateral region were severely impaired in learning this task, unlike patients with anterior temporal lesions, who learned the task as well as the normal control subjects (see Petrides, 1990, for details).

One of the conditional tasks administered to monkeys was analogous to the conditional task described above and was used with patients. The monkey was faced with two white Perspex boxes placed close to each other (see figure 15.5) (Petrides, 1985a). Inside each one of these boxes there was a light bulb that could be turned on or off by the experimenter. On each trial, one of the boxes, chosen according to a random sequence, was lit and the other remained unlit. The monkeys were rewarded if they opened the lit box when object A was shown and if they opened the unlit box when object B was shown. Thus, the monkeys had to learn to select between two visual nonspatial stimuli (i.e., the lit or the unlit box) based on a conditional rule. As can be seen in figure 15.5, monkeys with periarcuate lesions failed to learn this task within the limits of testing (1020 trials), in sharp contrast to the monkeys with middorsolateral prefrontal lesions who performed as well as the normal control animals. In another experiment, the monkeys were required to learn to perform one of two actions (grip a stick or touch a button) depending on the visual stimulus that was

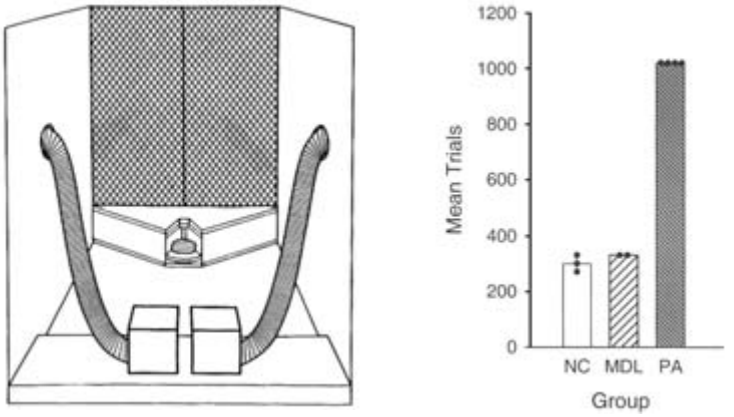


Figure 15.5
(*Left*) Schematic diagram of the experimental arrangement in the object conditional task administered to monkeys. On each trial, one of the two white Perspex boxes is lit and the other remains unlit by turning on a lightbulb that is inside them. One of two conditional stimuli is then presented in front of the opaque screen hiding the experimenter and the animal responds by pushing back one of the two boxes. The reward is delivered via the tubes that are attached to the boxes. (*Right*) Performance of animals with mid-dorsolateral frontal lesions (MDL), animals with periarcuate lesions (PA), and normal control animals (NC). Solid circles indicate the scores of individual animals in each group. Note that none of the animals with PA lesions was able to reach criterion within the limits of testing.

shown on any given trial (figure 15.6). Again, the monkeys with periarcuate lesions were severely impaired (Petrides, 1982). Further research showed that when the task requires the selection between distinct movements, the rostral dorsal area 6 is the critical one (Halsband & Passingham, 1982; Petrides, 1987). By contrast, when the animal has to select between distinct visual stimuli, the critical region is area 8 (Petrides, 1987).

The research with monkeys demonstrated that the caudal dorsolateral frontal cortex is a critical region for the performance of conditional associative tasks that require selection between different responses based on conditional rules. Control experiments showed that both patients and monkeys with caudal dorsolateral frontal lesions can perform the alternative responses, although they are impaired in selecting between these alternative responses based on conditional cues. Furthermore, these subjects (monkey and human) exhibit normal performance on many recognition and working-memory tasks, as well as on other tasks that require learning associations between stimuli. Thus, the problem on conditional associative tasks after caudal dorsolateral frontal lesions is not the result of a generalized impairment in learning; nor is it a problem in the performance or knowledge of the responses from which selections

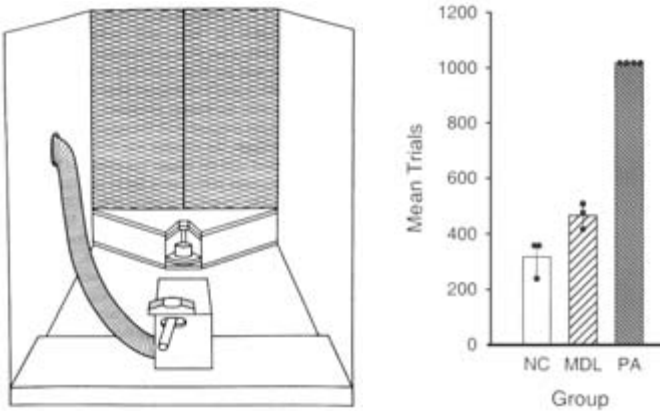


Figure 15.6
(*Left*) Schematic diagram of the experimental arrangement in the motor conditional task administered to monkeys. On each trial, one of two conditional stimuli is presented in front of the opaque screen hiding the experimenter and the animal responds by gripping the stick or touching the button. If the response is correct, reward is delivered via the tube that is attached to the manipulandum and the monkey pushes it back to receive the reward. (*Right*) Performance of animals with mid-dorsolateral frontal lesions (MDL), animals with periarcuate lesions (PA), and normal control animals (NC). Solid circles indicate the scores of individual animals in each group. Note that none of the animals with PA lesions was able to reach criterion within the limits of testing.

must be made. The impairment on the conditional tasks reflects a breakdown in the higher-order control of the selection between competing responses based on conditional rules.

As is the case with all cognitive and behavioral processing, conditional associative learning depends on several brain structures, not only the caudal dorsolateral frontal cortex. First, the learning of conditional associations would be expected to involve the medial temporal lobe limbic region (entorhinal/perirhinal cortex and the hippocampal system) that is critical for the learning of several different types of explicit associations between stimuli (see Eichenbaum, 1997). Consistent with this expectation, we found that patients with anterior temporal lobe lesions that did not involve the hippocampal system learned conditional associations as well as the normal control subjects, but those patients who had extensive damage to the hippocampal system were impaired in learning these associations (Petrides, 1985b, 1997). Findings from studies with both monkeys (Rupniak & Gaffan, 1987; Murray & Wise, 1996) and rats (Sziklas, Lebel, & Petrides, 1998) also agree that the hippocampal system is necessary for the learning of conditional associations.

Within the prefrontal cortex, Wang and colleagues (2000) found that bilateral injection of bicuculline (a GABAergic antagonist) into the ventrolateral prefrontal cortex

impaired the learning of novel visual-motor conditional associations, while leaving the performance of preoperatively learned ones intact. Infusions of bicuculline in the mid-dorsolateral prefrontal region did not affect the learning of novel visual-motor conditional associations or the performance of preoperatively learned ones, as would be expected from earlier work (e.g., Petrides, 1982; see figure 15.6). Furthermore, Bussey, Wise, and Murray (2001) found that massive lesions that involved not only the ventrolateral prefrontal cortex but also the orbitofrontal cortex as far as the medial orbital sulcus caused an impairment in the rapid learning (i.e., within a session of 50 trials) of novel visual-motor conditional associations, although these monkeys could learn the associations when trained gradually over several sessions. These findings show that ventrolateral and orbital prefrontal cortex play some role during the learning of novel visuomotor associations, especially when these associations are rapidly acquired, but it is not clear what this role might be. In order to be able to learn conditional associations within 50 trials, the monkeys in the Bussey, Wise, and Murray (2001) study had received considerable training preoperatively and developed a certain strategic approach to the task. The monkeys with the ventrolateral plus orbital frontal lesions had severe problems on the retention of these preoperatively learned strategies and also failed to learn other basic tasks such as matching-to-sample. Thus, in these monkeys, the difficulty in learning conditional visual-motor tasks appears to be reflecting more general cognitive impairments. This pattern of deficits contrasts sharply with the specific effects of periarcuate lesions. As pointed out above, the monkeys with periarcuate lesions fail both to acquire novel conditional associations and to perform preoperatively learned ones, while at the same time performing normally on matching-to-sample tasks, working memory tasks, and other associative learning tasks (see Petrides, 1987).

Functional Neuroimaging Studies

The dissociation between the caudal dorsolateral frontal cortex and the mid-dorsolateral frontal cortex that was demonstrated in lesion studies on monkeys has received confirmation from functional neuroimaging studies on normal human subjects. In an early study with positron emission tomography, we measured cerebral blood flow during the performance of a self-ordered task in which the subjects had to monitor their selections from a set of abstract designs, a visual conditional task in which the subjects had to select between the same set of visual stimuli based on conditional operations, and a visual matching control task. The same eight visual stimuli (abstract designs) were used in all three tasks, and these stimuli were presented in a different random arrangement on each trial. The subjects had to indicate their selections by pointing to particular stimuli. Thus, the only difference between the three tasks lay in their cognitive requirements. In the self-ordered task, which was directly

analogous to that previously used with patients and monkeys, the subjects were required to select a different design on each trial until all had been selected. The subjects were therefore required to monitor their selections as they performed the task. In the matching control task, the subjects were required to search and find a particular stimulus on each trial. This control task therefore involved perception of the same visual stimuli and the same searching behaviour as the self-ordered task, but did not require monitoring, that is, considering the earlier selections in relation to the current selections. In the conditional task, the subjects selected between the different designs based on arbitrary conditional relations between the different designs and different colored cues. These relations had been learned before scanning. In the conditional task, the designs and the searching among these designs were the same as in the self-ordered task, but since the design to be selected on each trial was determined by the color cue presented on that trial, correct performance did not require monitoring prior selections. Thus, whereas, on each trial, the self-ordered task required selection between the designs based on the monitoring of earlier selections, the conditional task required selection between the designs based on established conditional rules.

Performance of the self-ordered task, in comparison with either the matching control or the conditional task, resulted in significantly greater activity within the mid-dorsolateral prefrontal cortex (areas 46 and 9/46). There was no greater activity in this region when cerebral blood flow in the conditional task was compared with that of the matching control task, although there was now increased activity within the caudal dorsolateral frontal cortex in area 8. Note that area 8 is the very part of the caudal dorsolateral frontal region that studies in monkeys have shown to be critical for conditional operations applied to visual stimuli (see Petrides, 1987). Thus, the functional neuroimaging results were in agreement with the functional dissociation between mid-dorsolateral and caudal dorsolateral frontal cortex demonstrated in animal work. Other functional neuroimaging studies of conditional visuomotor learning have also yielded results consistent with the animal work (e.g., Deiber et al., 1997; Grafton et al., 1998).

The Rostral-Caudal Axis of Frontal Cortex Organization

The data presented above show that there is a rostral-caudal axis of frontal cortex organization (figures 15.7 and 15.8). The most caudal part of the frontal cortex, namely the motor region that occupies the precentral gyrus, consists of somatotopically organized motor representations (e.g., area 4, caudal area 6) that are primarily involved in fine motor control and sensorimotor transformations. It is important, however, to point out here that, whereas the caudal part of area 6, which lies on the precentral gyrus, is intricately involved in sensorimotor control, the rostral dorsal area 6, which in the monkey lines the caudal bank of the arcuate sulcus, can be considered to be

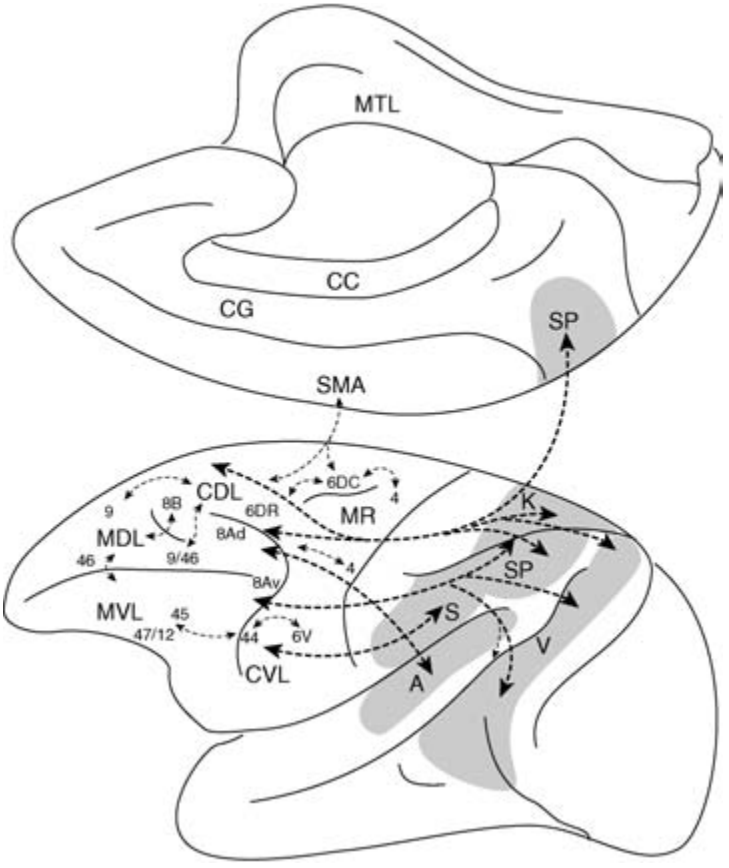


Figure 15.7

The rostral-caudal axis of lateral frontal cortex organization. Schematic diagram to illustrate some of the functional interactions postulated to underlie the caudal dorsolateral frontal region functional organization. Abbreviations: A, auditory processing in superior temporal gyrus; CC, corpus callosum; CG, cingulate gyrus; K, kinaesthetic processing in the superior parietal lobule; MTL, medial temporal lobe region; S, body-centered (i.e., somatocentric) amodal processing in rostral inferior parietal lobule; SMA, supplementary motor area; SP, spatial processing in posterior parietal cortex; V, visual processing.

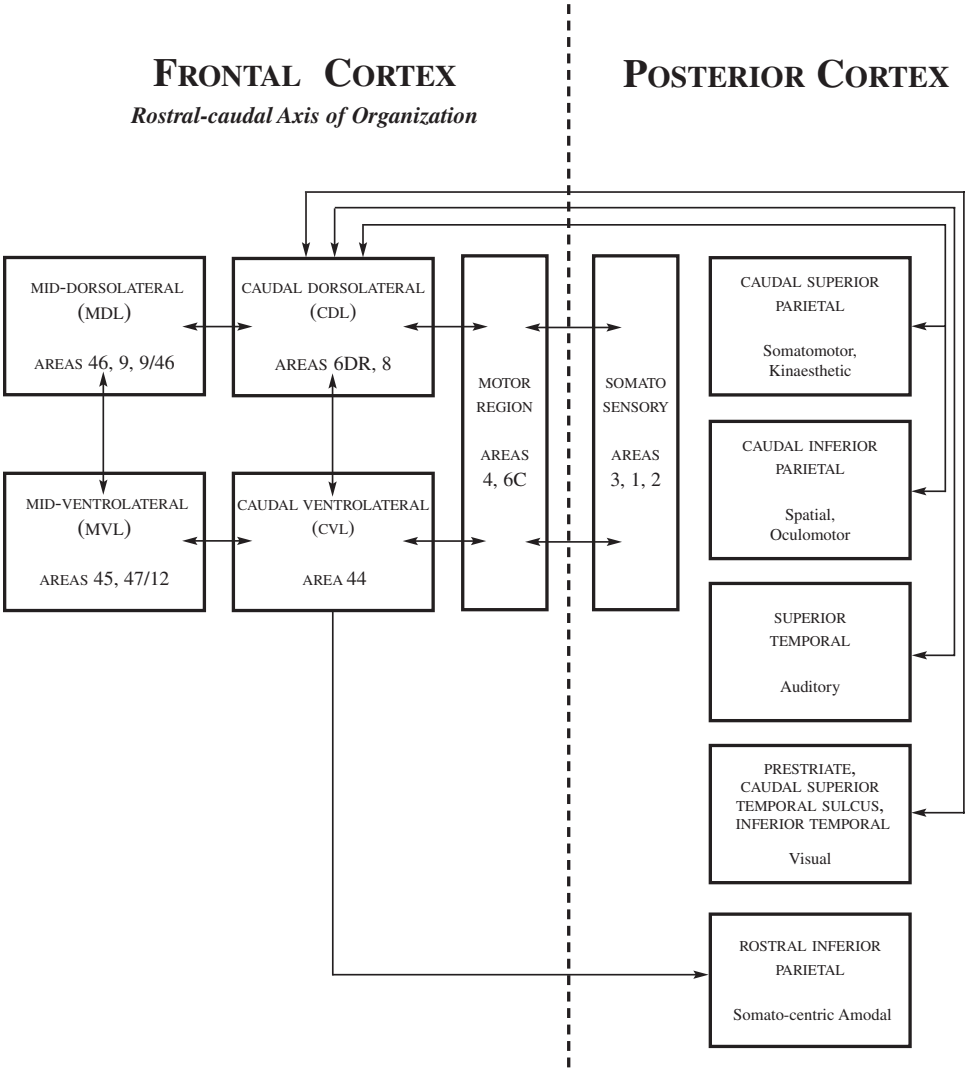


Figure 15.8
Flow diagram to illustrate some of the functional interactions postulated to underlie the rostral-caudal axis of organization of the frontal cortex.

part of the caudal dorsolateral frontal cortex in terms of function and certain connectional patterns (Picard & Strick, 2001; Rizzolatti & Luppino, 2001). Although rostral dorsal area 6 does play a role in motor control, this role lies in the selection between alternative motor responses based on conditional rules and, in this respect, it shares the function of the remaining caudal dorsolateral frontal cortex (see Petrides, 1987).

There is strong evidence that the caudal dorsolateral frontal region is critically involved in the selection between competing responses based on conditional operations (see Petrides, 1987). The caudal dorsolateral frontal region, however, comprises different parts that exhibit differences in their connections with posterior association cortex (figure 15.7). I have argued that all these sectors of the caudal lateral frontal region are involved in conditional selection, but that the conditional operations are applied to different types of information depending on the distinct connections of the different sectors of the caudal lateral frontal cortex with posterior association cortex. Figures 15.7 and 15.8 show some of the postulated functional interactions between different parts of the caudal lateral frontal region and posterior association cortex.

Rostral dorsal area 6 is strongly connected, locally, with motor areas, such as caudal area 6 and supplementary motor area (SMA). At the same time, rostral dorsal area 6 is connected with the superior parietal lobule and the caudal part of the inferior parietal lobule. Several investigations (Duffy and Burchfield, 1971; Sakata, Takaoka, Kawarasaki, & Shibutani, 1973; Mountcastle et al., 1975; Lacquaniti et al., 1995) have shown that the neurons of the cortex of the superior parietal lobule code the location of body parts (e.g., the arm) in a body-centered coordinate system. Thus, the rostral dorsal component of area 6 by virtue of its connections can play a major role in the selection between alternative competing motor acts based on conditional operations. There is strong evidence that this is the case (Petrides, 1982; Halsband and Passingham, 1982).

Area 8, which lies in front of the rostral dorsal area 6, is also involved in conditional operations, but in conditional operations that are applied to different types of information. Area 8Av, which is primarily linked with the visual system and the oculomotor and spatial processing in the caudal inferior parietal lobule (Mountcastle et al., 1975; Hyvärinen & Shelepin, 1979; Andersen & Gnadt, 1989; Goldberg & Segraves, 1989), can control the selection between alternative visual stimuli in the environment based on conditional rules (Petrides, 1987). By contrast, the more dorsal part of area 8, namely area 8Ad, which maintains strong connections with the caudal superior temporal gyrus is in a position to exercise conditional selection of auditory information.

Selection between different aspects of the current visual, auditory, and somatomotor environment based on conditional operations can be thought of as the conditional

allocation of attention to different competing stimuli in the environment. For instance, in a visual conditional task, the subject is required to select (i.e., allocate attention to) object X, if cue A, but allocate attention to object Y, if cue B. Similarly, in an auditory conditional task, among competing sounds in the environment, the subject has to allocate attention to (i.e., select) auditory stimulus X, if cue A, but focus attention to sound Y, if cue B. Thus, conditional associative learning and performance provides a means by which attention can be flexibly switched between different stimuli or responses in a given situation under different conditions based on prior learned arbitrary associations.

Although the role of the caudal ventrolateral frontal cortex has not been the object of investigation in accordance with the above concept, we suggest that it can probably be viewed by analogy with the caudal dorsolateral frontal areas. The caudal ventrolateral frontal cortex is linked strongly with the anterior portion of the inferior parietal lobule (see Petrides & Pandya, 2002). Neurons in the anterior inferior parietal lobule exhibit complex body-centered (i.e., somatocentric) responses (Hyvärinen & Shelepin, 1979; Leinonen, Hyvärinen, Nyman, & Linnankoski, 1979; Robinson & Burton, 1980; Taira et al., 1990). These responses, although centered on parts of the body or specific actions, may involve information arriving from other sensory modalities (e.g., visual). Thus, the caudal ventral frontal cortex (area 44 and adjacent rostral ventral area 6) are in a position to exercise high level control on certain aspects of action processing. We suggest that one aspect of this control may lie in the conditional selection of face and arm actions or the conditional allocation of attention to actions involving the body. The data showing that neurons in this frontal region respond during specific goal directed actions and appear to code action in high-level abstract terms rather than the movements per se (Rizzolatti, Camarda, & Fogassi, 1988; Rizzolatti, Fadiga, & Gallese, 1996; Murata, Fadiga, & Gallese, 1997) are consistent with the proposed notion.

Thus, there is a rostral-caudal axis of functional organization with the frontal cortex. The most caudal region is involved in direct sensorimotor mappings, whereas the caudal lateral frontal region is involved in higher order control processes that regulate the allocation of attention and therefore selection based on conditional operations. Further rostrally, the mid-lateral prefrontal region plays an even more abstract role in cognitive control. The mid-lateral prefrontal region is itself organized along a dorsal-ventral axis, with the dorsal mid-dorsolateral prefrontal cortex being involved in the monitoring of information in working memory and the mid-ventrolateral prefrontal region being involved in active judgements on information held in posterior cortical association regions that are necessary for active retrieval and encoding of information (see Petrides, 1994, 1996). In conclusion, the lateral prefrontal cortical areas are organized both along a rostral-caudal and a dorsal-ventral axis of organization.

Acknowledgments

This work was supported by grants from the National Sciences and Engineering Research Council of Canada and the Canadian Institutes of Health Research.

References

Andersen, R., & Gnadt, J. W. (1989). Role of posterior parietal cortex in saccadic eye movements. In R. Wurtz & M. Goldberg (Eds.), *The Neurobiology of Saccadic Eye Movements* (pp. 315–335). Amsterdam: Elsevier.

Bachevalier, J., & Mishkin, M. (1986). Visual recognition impairment follows ventromedial but not dorsolateral prefrontal lesions in monkeys. *Behavioural Brain Research*, *20*, 249–261.

Barbas, H., & Pandya, D. N. (1987). Architecture and frontal cortical connections of the premotor cortex (area 6) in the rhesus monkey. *Journal of Comparative Neurology*, *256*, 211–228.

Barbas, H., & Pandya, D. N. (1989). Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. *Journal of Comparative Neurology*, *286*, 353–375.

Brodmann, K. (1905). Beitraege zur histologischen Lokalisation der Grosshirnrinde. III. Mitteilung: Die Rindenfelder der niederen Affen. *Journal of Psychological Neurology*, *4*, 177–226.

Brodmann, K. (1908). Beitraege zur histologischen Lokalisation der Grosshirnrinde. V1. Mitteilung: Die Cortexgliederung des Menschen. *Journal of Psychological Neurology*, *10*, 231–246.

Brodmann, K. (1909). *Vergleichende Localisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Barth: Leipzig.

Bussey, T. J., Wise, S. P., & Murray, E. A. (2001). The role of ventral and orbital prefrontal cortex in conditional visuomotor learning and strategy use in rhesus monkeys (*Macaca mulatta*). *Behavioral Neuroscience*, *115*, 971–982.

Deiber, M. P., Wise, S. P., Honda, M., Catalan, M. J., Grafman, J., & Hallett, M. (1997). Frontal and parietal networks for conditional motor learning: A positron emission tomography study. *Journal of Neurophysiology*, *78*, 977–991.

Duffy, F., & Burchfield, J. (1971). Somatosensory system: Organizational hierarchy from single units in monkey area 5. *Science*, *172*, 273–275.

Economo, C., & Koskinas, G. (1925). *Die Cytoarchitektonik der Hirnrinde des erwachsenen Menschen* (The cytoarchitectonics of the human cerebral cortex). Berlin: Springer.

Eichenbaum, H. (1997). Declarative memory: Insights from cognitive neurobiology. *Annual Review of Psychology*, *48*, 547–572.

Geyer, S., Ledberg, A., Schleicher, A., Kinomura, S., Schormann, T., Burgel, U., Klingberg, T., Larsson, J., Zilles, K., & Roland, P. E. (1996). Two different areas within the primary motor cortex of man. *Nature*, *382*, 805–807.

Goldberg, M. E., & Segraves, M. A. (1989). The visual and frontal cortices. *Review of Oculomotor Research*, *3*, 283–313.

Goldman-Rakic, P. S. (1996). Architecture of the prefrontal cortex and the central executive. *Annals of the New York Academy of Sciences*, *769*, 71–83.

Grafton, S. T., Fagg, A. H., & Arbib, M. A. (1998). Dorsal premotor cortex and conditional movement selection: A PET functional mapping study. *Journal of Neurophysiology*, *79*, 1092–1097.

Grafton, S. T., Hari, R., & Selenius, S. (2000). The human motor system. In A. W. Toga & J. C. Mazziotta (Eds.), *Brain Mapping: The Systems* (pp. 331–363). San Diego: Academic Press.

Halsband, U., & Passingham, R. (1982). The role of premotor and parietal cortex in the direction of action. *Brain Research*, *240*, 368–372.

He, S. Q., Dum, R. P., & Strick, P. L. (1993). Topographic organization of corticospinal projections from the frontal lobe: Motor areas on the lateral surface of the hemisphere. *Journal of Neuroscience*, *13*, 952–980.

Hyvärinen, J., & Shelepin, Y. (1979). Distribution of visual and somatic functions in the parietal associative areas 7 of the monkey. *Brain Research*, *169*, 561–564.

Lacquaniti, F., Guigon, E., Bianchi, L., Ferraina, S., & Caminiti, R. (1995). Representing spatial information for limb movement: Role of area 5 in the monkey. *Cerebral Cortex*, 5, 391–409.

Leinonen, L., Hyvårinen, J., Nyman, G., & Linnankoski, I. (1979). Functional properties of neurons in lateral part of associative area 7 in awake monkeys. *Experimental Brain Research*, *34*, 299–320.

Luria, A. R. (1969). Frontal-lobe syndromes. In P. J. Vinken & G. W. Bruyn (Eds.), *Handbook of Clinical Neurology* (Vol. 2, pp. 725–757). Amsterdam: North Holland.

Mountcastle, V. B., Lynch, J. C., Georgopoulos, A., Sakata, H., & Acuna, C. (1975). Posterior parietal association cortex of the monkey: Command functions for operations within extrapersonal space. *Journal of Neurophysiology*, *38*, 871–908.

Murata, A., Fadiga, L., & Fogassi, L. (1997). Object representation in the ventral premotor cortex (area F5) of the monkey. *Journal of Neurophysiology*, *78*, 2226–2230.

Murray, E. A., & Wise, S. P. (1996). Role of the hippocampus plus subjacent cortex but not amygdala in visuomotor conditional learning in rhesus monkeys. *Behavioral Neuroscience*, *110*, 1261–1270.

Penfield, W., & Boldrey, E. (1937). Somatic and motor representation in the cerebral cortex of man as studied by electrical stimulation. *Brain*, *60*, 389–443.

Petrides, M. (1982). Motor conditional associative-learning after selective prefrontal lesions in the monkey. *Behavioral Brain Research*, *5*, 407–413.

Petrides, M. (1985a). Deficits in non-spatial conditional associative learning after periarcuate lesions in the monkey. *Behavioral Brain Research*, *16*, 95–101.

Petrides, M. (1985b). Deficits on conditional associative-learning tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia*, *23*, 601–614.

Petrides, M. (1986). The effect of periarcuate lesions in the monkey on the performance of symmetrically and asymmetrically reinforced visual and auditory go, no-go tasks. *Journal of Neuroscience*, *6*, 2054–2063.

Petrides, M. (1987). Conditional learning and the primate frontal cortex. In E. Perecman (Ed.), *The Frontal Lobes Revisited* (pp. 91–108). New York: IRBN Press.

Petrides, M. (1990). Nonspatial conditional learning impaired in patients with unilateral frontal but not unilateral temporal lobe excisions. *Neuropsychologia*, *28*, 137–149.

Petrides, M. (1991). Monitoring of selections of visual stimuli and the primate frontal cortex. *Proceedings of the Royal Society of London, Series B*, *246*, 293–298.

Petrides, M. (1994). Frontal lobes and working memory: Evidence from investigations of the effects of cortical excisions in nonhuman primates. In F. Boller & J. Grafman (Eds.), *Handbook of Neuropsychology* (Vol. 9, pp. 59–82). Amsterdam: Elsevier.

Petrides, M. (1995). Impairments on nonspatial self-ordered and externally ordered working memory tasks after lesions of the mid-dorsal part of the lateral frontal cortex in the monkey. *Journal of Neuroscience*, *15*, 359–375.

Petrides, M. (1996). Specialized systems for the processing of mnemonic information within the primate frontal cortex. *Philosophical Transactions of the Royal Society, London, Series B*, *351*, 1455–1462.

Petrides, M. (1997). Visuo-motor conditional associative learning after frontal and temporal lesions in the human brain. *Neuropsychologia*, *35*, 989–997.

Petrides, M. (2000a). Frontal lobes and memory. In F. Boller & J. Grafman (Eds.), *Handbook of Neuropsychology* (2nd ed., Vol. 2, pp. 67–84). Amsterdam: Elsevier.

Petrides, M. (2000b). Mapping prefrontal cortical systems for the control of cognition. In A. W. Toga & J. C. Mazziotta (Eds.), *Brain Mapping: The Systems* (pp. 159–176). San Diego: Academic Press.

Petrides, M. (2000c). Middorsolateral and midventrolateral prefrontal cortex: Two levels of executive control for the processing of mnemonic information. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII* pp. 535–548. Cambridge, MA: The MIT Press.

Petrides, M. (2000d). Dissociable roles of mid-dorsolateral and anterior inferotemporal cortex in visual working memory. *Journal of Neuroscience*, *20*, 7496–7503.

Petrides, M., Alivisatos, B., Evans, A. C., & Meyer, E. (1993). Dissociation of human mid-dorsolateral from posterior dorsolateral frontal cortex in memory processing. *Proceedings of the National Academy of Sciences*, *90*, 873–877.

Petrides, M., & Milner, B. (1982). Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia*, *20*, 249–262.

Petrides, M., & Pandya, D. N. (1994). Comparative architectonic analysis of the human and the macaque frontal cortex. In F. Boller & J. Grafman (Eds.), *Handbook of Neuropsychology* (Vol. 9, pp. 17–58). Amsterdam: Elsevier.

Petrides, M., & Pandya, D. N. (2002). Association pathways of the prefrontal cortex and functional considerations. In D. T. Stuss & R. T. Knight (Eds.), *Principles of Frontal Lobe Function* (pp. 31–50). Oxford: Oxford University Press.

Picard, N., & Strick, P. L. (2001). Imaging the premotor areas. *Current Opinions in Neurobiology*, *11*, 663–672.

Rizzolatti, G., Camarda, R., & Fogassi, L. (1988). Functional organization of inferior area 6 in the macaque monkey: II. Area F5 and the control of distal movements. *Experimental Brain Research*, *71*, 491–507.

Rizzolatti, G., Fadiga, L., & Gallese, V. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*, 131–141.

Rizzolatti, G., & Luppino, G. (2001). The cortical motor system. *Neuron, 31*, 889–901.

Rizzolatti, G., Luppino, G., & Matelli, M. (1998). The organization of the motor system: New concepts. *Electroencephalographic Clinical Neurophysiology*, *106*, 283–296.

Robbins, T. W. (1996). Dissociating executive functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London B Biological Science*, *351*(1346), 1463–1470.

Robinson, C. J., & Burton, H. (1980). Organization of somatosensory receptive fields in cortical areas 7b, retroinsula, postauditory, granular insula of *M. fascicularis*. *Journal of Comparative Neurology*, *192*, 69–92.

Rupniak, N. M. J., & Gaffan, D. (1987). Monkey hippocampus and learning about spatially-directed movements. *Journal of Neuroscience*, *7*, 2331–2337.

Sakata, H., Takaoka, Y., Kawarasaki, A., & Shibutani, H. (1973). Somatosensory properties of neurons in the superior parietal cortex (area 5) of the rhesus monkey. *Brain Research*, *64*, 85–102.

Shallice, T., & Burgess, P. (1996). The domain of supervisory processes and temporal organization of behaviour. *Philosophical Transactions of the Royal Society of London B Biological Science*, *351*(1346), 1405–1411.

Stuss, D. T., Benson, D. F., Clermont, R., Della Malva, C. L., Kaplan, E. F., & Weir, W. S. (1986). Language functioning after bilateral prefrontal leukotomy. *Brain and Language*, *28*(1), 66–70.

Sziklas, V., Lebel, S., & Petrides, M. (1998). Conditional associative learning and the hippocampal system. *Hippocampus*, *8*, 131–137.

Taira, M., Mine, S., Georgopoulos, A. P., Murata, A., & Tanaka, Y. (1990). Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Experimental Brain Research*, *83*, 29–36.

Walker, A. E. (1940). A cytoarchitectural study of the prefrontal area of the macaque monkey. *Journal of Comparative Neurology*, *73*, 59–86.

Wang, M., Zhang, H., & Li, B.-M. (2000). Deficit in conditional visuomotor learning by local infusion of bicuculline into the ventral prefrontal cortex in monkeys. *European Journal of Neuroscience*, *12*, 3787–3796.

16 Primate Anterior Cingulate Cortex and Adaptation of Behavior

Céline Amiez, Jean-Paul Joseph, and Emmanuel Procyk

By just looking around and observing humans coping with the complex environments they have constructed, one can see that our species is characterized by a fantastic ability to adapt. Humans spend most of their life solving problems. Problem solving, reasoning, and other complex activities necessitate executive functions that have greatly evolved in primate to reach yet incomparable levels in humans. These activities require complex successions of processes like attending, deciding and choosing, acting, detecting mistakes, and changing responses.

The neurobiological bases of executive functions depend mostly on the integrity of the frontal lobe, and of its relationships with other cortical and subcortical structures. The anterior cingulate cortex (ACC), subdivision of the frontal lobe, is particularly linked to voluntary behavior and adaptive functions. This explains in part why malfunctioning of this structure is devastating, and appears in many psychiatric disorders. Here we review findings from human and monkey ACC anatomy, neuropsychology, and physiology in an attempt to highlight their commonalities, and we detail how data from these fields participate in the design of complex models of primate executive functions. Although this review focuses on this one cortical region, we do not wish to leave a purely localisationist impression. The ACC gets its specificities from its position within several functional networks, which *in fine* are the key to understanding the neurobiological substrate of cognitive functions. Nevertheless, as a first step we need to know about the specificities of each node of the networks.

Anatomy

The cingulate cortex is located on the medial surface of the hemispheres, covering the entire length of the corpus callosum. In both humans and monkeys, it is divided into an anterior agranular region (Brodmann's areas 24a/b/c, 25, 32) and a posterior granular region (BA 23a/b/c, 29, 30, 31). In humans, areas 32′ and 33 are also part of the anterior region. The cortex of the cingulate sulcus is distinguished from the cortex of the cingulate gyrus and comprises areas 23c, 24c, 32, and 32′ in humans and areas

24c, 23c, and 6c in monkeys. The gyrus contains areas 23a, 23b, 24a, 24b, 25, and 32 in both humans and monkeys, and area 33 in humans (figure 16.1).

The caudal limitation of the anterior cingulate cortex (ACC) is the dorsal extremity of the central sulcus. Consequently ACC is part of the medial frontal lobe. In both humans and monkeys, ACC contains areas 24a, 24b, 24c, 25, and 32. Areas 32′ and 33, which are unique to humans, are also part of the ACC (figure 16.1). In both monkeys and humans, area 24c is buried into the rostral part of the cingulate sulcus and borders area 24b. Area 24c covers the ventral and, depending on authors, the dorsal banks of the cingulate sulcus. Area 32 which is located rostral to area 24b is considered as transition between cingulate and prefrontal cortex (Carmichael & Price, 1994; Vogt, 1993; Vogt, Nimchinsky, Vogt, & Hof, 1995). In both humans and monkeys, ventral and dorsal regions are heavily interconnected, nevertheless they belong to two distinct functional systems.

Studies of connectivity, using injections of anterograde and retrograde axonal tracers, have demonstrated that (1) ACC subdivisions are strongly interconnected. For instance, 24c is heavily interconnected with areas 24b, 23b, and 23c, moderately interconnected with areas 24a and 23a, and has few connections with subgenual area 25, area 32 (Morecraft & Van Hoesen, 1998) and (2) ACC, as a whole, is connected with various cortical and subcortical structures, and thus takes part in several domains of information processing (figure 16.2).

ACC has strong connections with the prefrontal cortex (Barbas & Mesulam, 1985; Bates & Goldman-Rakic, 1993; Lu, Preston, & Strick, 1994; McGuire, Bates, & Goldman-Rakic, 1991; Morecraft, Genla, & Mesulam, 1993; Pandya, Dye, & Buttors, 1971; Pandya & Vignolo, 1971; Selemon & Goldman-Rakic, 1988), with the limbic cortex (Barbas & Pandya, 1989; Carmichael & Price, 1995; Insausti, Amaral, & Cowan, 1987) and with motor and premotor cortices. Subdivisions of the sulcus that connect to the primary motor cortex are named the cingulate motor areas (CMA). Dum and Strick (1991, 1993) have described a rostral area—CMAr—that occupies both the dorsal and ventral banks of the sulcus, and a caudal area subdivided in CMAv and CMAd, respectively located in the ventral and dorsal banks (Dum & Strick, 1991, 1993). CMAr, CMAv, and CMAd correspond to subfields of areas 24c, 23c, and 6c (Dum & Strick, 1991; Luppino, Matelli, Camarda, & Rizzolatti, 1991). Each of these three fields is somatotopically organized and contains a face, leg, and hand representations (Dum & Strick, 1991; He, Dum, & Strick, 1995; Luppino et al., 1991). In addition, ACC is connected with cortical areas involved in motor control: supplementary motor area (SMA), pre–SMA, lateral premotor cortex, and oculomotor fields (FEF and SEF) (Huerta, Krubitzer, & Kaas, 1987; Luppino et al., 1993, 1994; Luppino, Rozzi, Calzavara, & Matelli, 2003; P. K. McGuire et al., 1991; Morecraft & Van Hoesen, 1992, 1993; Wang, Shima, Sawamura, & Tanji, 2001). Other main cortical connections are made with parietal, temporal, and insular cortices (Cipolloni & Pandya, 1999; Marconi et al., 2001).

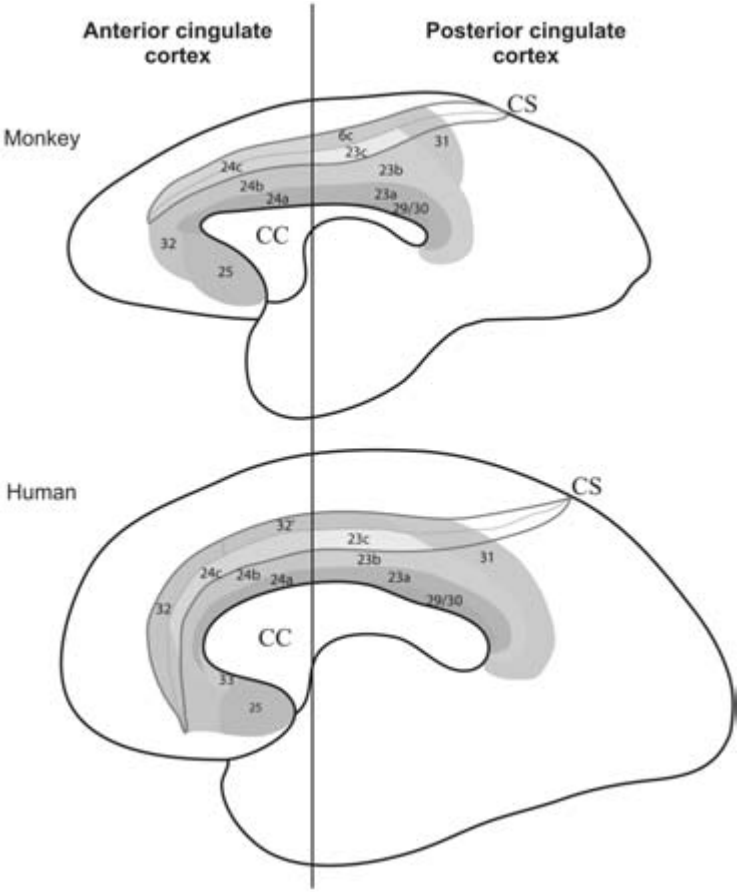


Figure 16.1

Schematic cytoarchitectonic subdivisions of the cingulate cortex in monkey and human. Homolog areas in monkey and human are represented with the same color. The cingulate sulcus is represented open, with its fundus delineating the ventral and the dorsal banks. The vertical line separates anterior from posterior cingulate cortex. The anterior cingulate cortex includes areas 24, 25, and 32. Posterior cingulate cortex includes areas 23, 29, 30, and 31. Abbreviations: CS, cingulate sulcus; CC, corpus callosum (see Carmichael & Price, 1994; He, Dum, & Strick, 1995; Matelli, Luppino, & Rizzolatti, 1991; Vogt, 1993; Vogt et al., 1987).

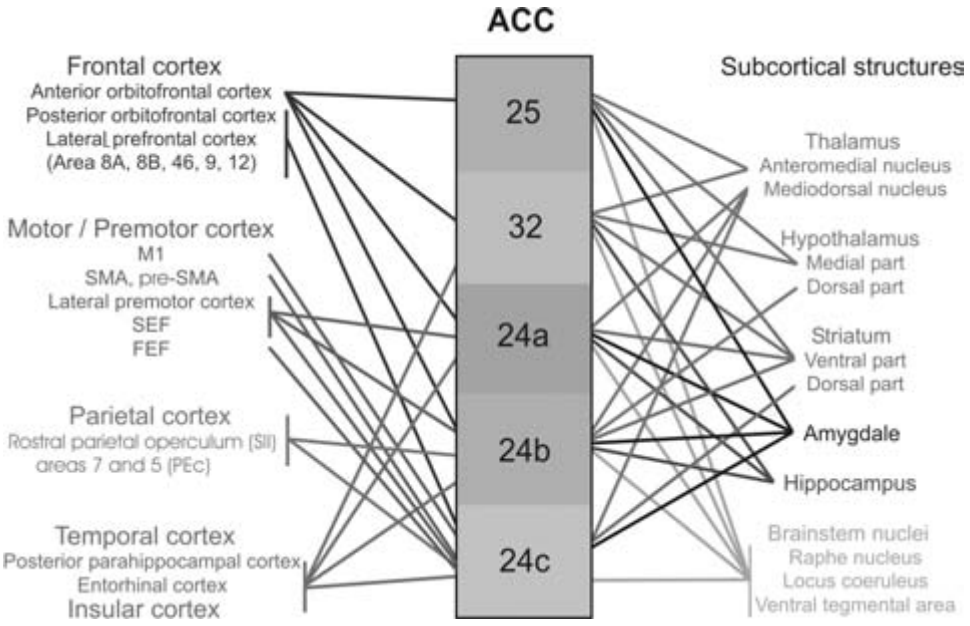


Figure 16.2
Summary of ACC's connectivity with frontal, temporal, motor, and premotor cortices and with subcortical structures. Abbreviations: M1, primary motor cortex; SMA, supplementary motor area; FEF, frontal eye field; SEF, supplementary eye field. See text for references.

Finally, ACC is connected with the thalamus (Arikuni, Sakai, & Kubota, 1983; Siwek & Pandya, 1991; Vogt et al., 1987), striatum (Arikuni & Kubota, 1986; Kunishio & Haber, 1994), hypothalamus (Rosene & Van Hoesen, 1977) and brainstem (Crino, Morrison, & Hof, 1993). Serotoninergic (from the dorsal and medial raphe nucleus), noradrenergic (locus coeruleus), and dopaminergic afferents (ventral tegmental area) are scattered in every cortical layers of ACC areas. Noradrenergic and dopaminergic afferents are more dense than serotoninergic afferents (Crino, Morrison, & Hof, 1993).

In conclusion, the ventral ACC (i.e., area 24a, 24b, 25, and 32) is primarily connected with limbic structures whereas the dorsal region (i.e., area 24c) is more interconnected with prefrontal, motor, and premotor cortices. The separation between ventral and dorsal, or anterior and posterior ACC led to the definition of affective and cognitive (or executive) divisions (Bush, Luu, & Posner, 2000; Vogt et al., 1992). More specifically, the cortex lying in the sulcus (area 24c) has a particular role in executive functions (see below). In the context of behavioral adaptation, it is interesting to note that 24c is positioned as a node between structures devoted to reward processing and learning (including ventral striatum, dopaminergic afferences, and orbitofrontal

cortex), and a network devoted to working memory and procedural learning (prefrontal cortex, preSMA, dorsal striatum). These data help define a structural substrate to known interactions between emotion, cognition, and action (Paus, 2001). Functional hypotheses gained a lot from anatomical observations, but also from functional studies in monkey and humans that we will now detail.

Electrophysiological and Lesion Studies in Monkey

ACC lesions, reversible deactivations, and electrophysiological studies in monkeys have been made in various emotional, motor, and cognitive contexts. Yet, most of the recent studies focused on executive processes in the framework of reward-based behaviors. Executive processes refer to a set of cognitive skills responsible for the control of complex behavior in new or problematic situations. Funahashi defined these processes as "a product of the co-ordinated operation of various processes to accomplish a particular goal in a flexible manner." These processes include selection, sequencing complex actions, and error monitoring (Funahashi, 2001).

Voluntary Selection and Reward

Lesions of the cingulate gyrus (including areas 24, 25, and 32) excluding the dorsal bank of the cingulate sulcus impairs searching: monkeys are unable to temporally and spatially organize their behavior to search for food (Stern & Passingham, 1994). This type of lesion induces impairments in internally generated but not externally triggered movements unless lesions include the medial frontal gyrus (including areas SMA, preSMA) (Chen et al., 1995; Thaler et al., 1995). The dissociation between self-generated and externally driven movements was observed at the neuronal level more often in CMAr than in posterior areas (Shima et al., 1991). CMAr, CMAd, and CMAv have different functional characteristics (Cadoret & Smith, 1997; Picard & Strick, 1997).

A key to understand rostral ACC function might be movement selection based on reward. Shima and Tanji trained monkeys to select voluntarily and shift movements based on changes in reward value (Shima & Tanji, 1998). After muscimol injections in CMAr, so as to temporarily inactivate the area, the animals failed to change movement even when the reward was a lot reduced. However, response selection was not impaired when guided by an external auditory stimulus. The authors concluded that CMAr has a key role in voluntary movement selection based on reward, by establishing a link between motor and reward information. Unit recordings during this task revealed cells with increased discharge between the occurrence (and detection) of the reduced reward and initiation of a newly selected movement. The role of ACC in reward-based selection is also supported by Hadland and colleagues, who tested monkeys with bilateral ACC lesions in a reward-conditional response selection task

and in a visual discrimination-learning task (Hadland et al., 2003). ACC lesions impaired performance only in the reward-based selection task.

Adaptation of Behavior

As Shima and Tanji showed, the anterior portion of the cingulate sulcus has a key role for adaptation based on reward delivery. This phenomenon was further studied using a sequential problem-solving task (Procyk, Tanaka, & Joseph, 2000). The animal had to search by trial and error the correct sequence for touching three fixed spatial targets to get a reward. After discovery, the sequence was repeated four times. Then the correct order was changed, requiring a new search. A majority of ACC neurons related to sequence execution had activity specific to the trial and error periods, when behavioral flexibility was required (nonroutine). Activity in those neurons ceased as soon as the animal had accumulated enough information to infer the solution, but had not yet tested it. During repetition (routine), other task-related neurons encoded the serial order of movements (figure 16.3). Further analysis indicated that serial order activity reflected the rank of movements in reference to the end of the sequence, that is, the distance to reward delivery (Procyk & Joseph, 2001). A recent study neatly confirmed these results (Shidara & Richmond, 2002). Thus, ACC neurons participate in the evaluation of reward proximity (during sequences) and/or of the certainty to get a reward (during trial and error).

Task-related cells in the ACC are timely related to stimulus onset, arm movement, or reward delivery, but all can be modulated by the size of the expected reward (figure 16.3). In fact, reversible inactivations of the ACC impair strategic behavior during a choice task based on reward size. In successive trials, two fixed unknown visual stimuli (A and B) were presented on a touch screen, and the animal had to touch one of the two to get a liquid reward. Stimulus A provided 1.2 ml of liquid with the probability of 0.7 or 0.4 ml with the probability of 0.3. Stimulus B provided the same rewards, but with the opposite probabilities. The appropriate strategy to maximize the reward was to search for stimulus A (optimal stimulus) and to maintain this choice in successive trials. Inactivation of ACC with muscimol impaired the discovery of the optimal stimulus. The animals tended to maintain their first choice regardless of the amount of reward obtained. They were unable to use the reward obtained in one trial as a cue for the next trial(s). During this choice task, reward information imbedded in neural activity corresponded to the average reward given during trials by choosing the optimal stimulus. This complex reward information appears to be crucial to the adaptation of behavior to novel and/or changing situations.

Error

ACC lesioned monkeys have difficulty in correcting their behavior after erroneous trials (Rushworth, Hadland, Gaffan, & Passingham, 2003). This may reflect a role for

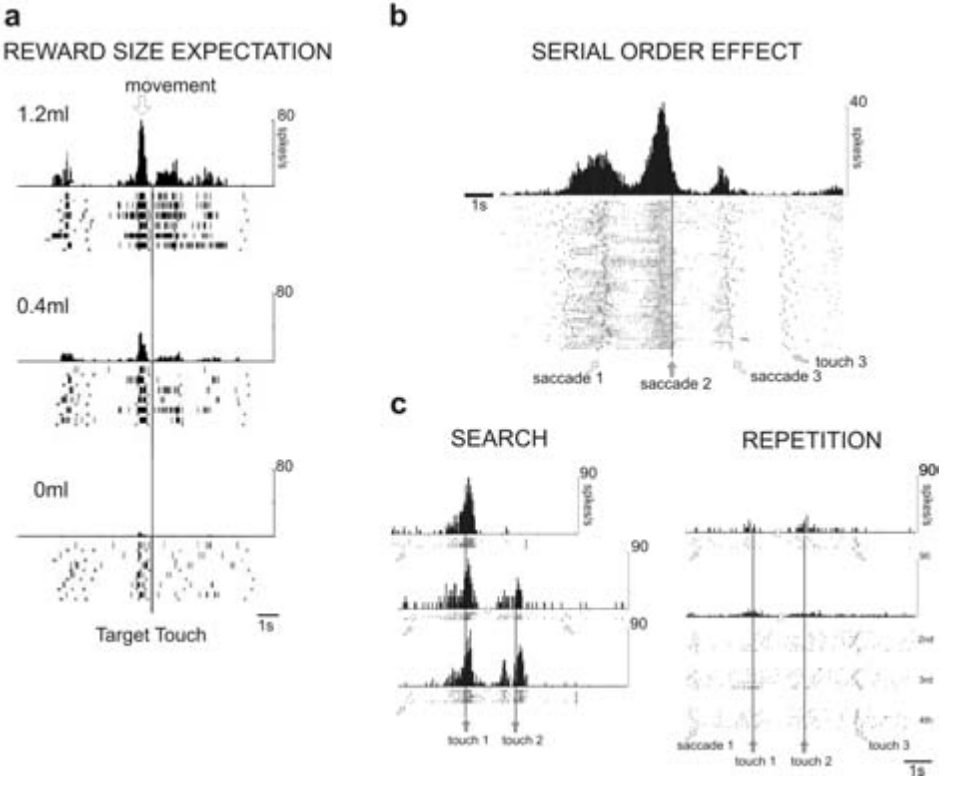


Figure 16.3

Activity of ACC cells. (*a*) Arm movement-related cell modulated by the size of expected reward. The activity was recorded during a stimulus-reward association task where well-known stimuli, associated with particular amount of reward (indicated on the left), had to be touched. (*b*) Cell showing variation of activity depending on the serial order of movements in sequences of three oculomotor saccades. (*c*) Cell with high activity in search trials (left) and low activity in repetition trials (right) during a problem solving task (see text for details). (Adapted from Procyk & Joseph, 2001; Procyk, Tanaka, & Joseph, 2000.)

ACC in error monitoring although other interpretations—insensitivity to change in reinforcer value or deficit in attention to actions' outcomes—are possible. To date, few experiments have described error signals in monkeys ACC (Brooks, 1986; Niki & Watanabe, 1976; Procyk & Joseph, 2001; Schall, Stuphorn, & Brown, 2002). The most detailed account reports error-related potentials generated by the ACC and particularly during insightful stages of learning (Brooks, 1986). In other words, there was no error signal during the very first phase of learning when monkeys had no clue about reward and response contingencies. Single unit recordings also report error-related

activity in CMAr. Recent work suggest a role for cingulate error-related activity in evaluation (Ito, Stuphorn, Brown, & Schall, 2003), and indeed, these activities seem to be more than an alert signal (Procyk et al., 2001; Procyk, Amiez, and Joseph, submitted). As we will see with human data, error detection might be a key feature of ACC function.

Experimental and Clinical Studies in Humans

It is sometimes difficult to make straight comparisons between monkey and human data, especially because of reward-oriented behaviors inherent to animal studies. This is why there are still protocols or even concepts that are used only in one or the other field of research.

In humans, ACC lesions originate from anterior cerebral artery damage or therapeutic cingulotomy in patients treated for severe psychiatric disorders, epilepsy, or intractable pain. It is rare when such lesions are limited to the cingulate cortex. They often include orbitofrontal cortex, supplementary motor areas, and subcortical structures. Nevertheless, valuable cases appear in the literature with limited lesion following small infarcts or tumour removal (e.g., Gaymard et al., 1998; Turken & Swick, 1999).

Brain imaging and event-related potentials (ERP) studies in humans have provided a massive amount of data and created most of the debate on ACC function(s). Two general observations have been made: (1) ACC activations appear during execution of a wide range of motor, cognitive, and affect-related (pain, emotion, etc.) tasks; (2) different ACC subdivisions are delineated, confirming anatomical and electrophysiological data in monkeys.

If one wants to summarize the extensive literature devoted to human ACC function, it is reasonable to focus on data pointing at the adaptation of behavior in challenging situations. Here, keywords are free selection, attention, conflict, error, and reward, and most of the data will concern the cognitive and motor parts of ACC.

Voluntary Selection

Bilateral lesions of ACC induce akynetic mutism, which leads to a dramatic reduction of voluntary initiation of verbal and motor responses (Devinksy & Luciano, 1993). Turken and Swick have studied the performance of a patient with focal and rather caudal ACC lesion in cognitive and simple tasks. The patient did not show specific cognitive impairment. On the other hand, he was more impaired for manual than for vocal responses, suggesting a functional motor specialization in ACC (Turken & Swick, 1999).

Several studies have focused on free selection of responses. Free selection, or self-initiation of movement, is also a condition during which monkey's CMAr neurons

respond specifically (Shima et al., 1991). In humans, verbal or motor fluency performance, as well as learning tasks, induces lateral prefrontal and ACC activation (Blakemore, Rees, & Frith, 1998; Deiber et al., 1991; Frith, Friston, Liddle, & Frackowiak, 1991; Paus, Petrides, Evans, & Meyer, 1993; Phelps, Hyder, Blamire, & Shulman, 1997). In general, complex cognitive tasks like the Tower of London or the Wisconsin card sorting, and various other cognitively challenging tasks induce lateral prefrontal and ACC activations (Duncan & Owen, 2000). Overall, activation of these areas appear for new situations, or learning, and disappear after routinization (Jenkins et al., 1994; Jueptner et al., 1997a, 1997b; Raichle et al., 1994), a phenomenon that may be compared to differences in monkey's cingulate unit activity between nonroutine and routine situations (Procyk et al., 2000).

Attention for Action and Conflict Monitoring

Free selection of movement involves particular attentional phenomenon devoted to motor control and outcome evaluation. ACC lesions induce attentional deficits. Patients showing focal lesion of ACC present impairments in response intention and focused attention (Cohen et al., 1999a, 1999b).

Remarkably, Jueptner and colleagues showed an increase of dorsolateral prefrontal and anterior cingulate cortices activations when subjects attend to as opposed to when they do not attend to the execution of overlearned motor sequences (Jueptner et al., 1997a, 1997b). Related studies showed that implicit sequence learning, as tested by serial reaction time (SRT), does not evoke ACC activation except for a few subjects that explicitly detect the presence of sequential patterns (Grafton, Hazeltine, & Ivry, 1995).

Simple designed—and so maybe more specific—tasks can also involve the ACC. For instance, a small region of dorsal ACC is activated during conventional delay tasks (Petit, Courtney, Ungerleider, & Haxby, 1998). Besides, the most famous results about ACC concern the performance of Stroop-like tasks. In the original version of the task, subjects must name the ink color of words, which are color names. During testing, incongruent trials ("blue" written in red) induce a decrease in response time and frequent errors, supposedly because of a conflict created by coactivation of an automatic response (reading the word) and the response imposed by the instruction (say, the ink color). When contrasting brain activity during incongruent and congruent (e.g., "blue" written in blue) trials, early experiments found activations of a mid-rostral subdivision of ACC (Bench et al., 1993; George et al., 1994; Pardo, Pardo, Janer, & Raichle, 1990).

Those results led to several hypotheses about a central role of rostral ACC as an "anterior attentional system" involved in attention to action (Posner & DiGirolamo, 1998; Posner & Petersen, 1990), or as a conflict monitor (Carter et al., 1998) (see IV-Models).

Error

Studies on learning and adaptive behavior reveal ACC activations associated to error detection. Indeed, using ERP, anterior cingulate was found to generate a negative potential (error-related negativity: ERN, or N_E) at the time of response error or of an external error feedback (Dehaene, Posner, & Tucker, 1994; Frankenstein, Richter, McIntyre, & Remy, 2001).

Error potential and error-related unit activities are also observable in monkeys' ACC (Brooks, 1986; Procyk & Joseph, 2001). With fMRI, an activation of ACC appears for incorrect trials (Ullsperger & von Cramon, 2003). Further, after focal ACC lesion, one patient showed altered ERN and deficits in the correction of erroneous responses (Swick & Turken, 2002). The ERN is at the center of the debate on the function of the cognitive part of ACC. It might relate to error detection itself, during motor execution or after perception of external feedback (see Holroyd & Coles, 2002, for review), but whether it is specific to error, and whether ACC is the sole origin, is still a matter of debate (Vidal, Hasbroucq, Grapperon, & Bonnet, 2000). Indeed, some authors observed a negative potential also in correct trials (CRN) and thus cast doubts on the nature of the ERN (see also below, Gehring & Willoughby, 2002).

Reward

Adaptation of behavior depends on not only negative but also positive feedback. Indeed, learning theories are deeply based on positive reinforcement, or reward.

Patients exhibiting ACC lesion coupled with orbitofrontal cortex lesion, always favored choices bringing large immediate reward and a loss in the long run, rather than choices leading to small immediate reward but a gain in the long run (assessed by the Iowa gambling task; Bechara et al., 1994). This needs to be compared with work in monkeys emphasizing a central function of ACC in behavioral selections based on food reward. Bridging animal and human studies, Bush and colleagues observed activation of human ACC in reward based tasks similar to those used in monkey studies (Bush et al., 2002). Of course, for humans the reward is in dollars; nevertheless, the results of Bush and colleagues are very similar to those obtained in electrophysiological studies. Interestingly, this study stressed the fact that the same region of ACC (but different cell populations) might encode error and reward. In the same idea, Gehring and Willoughby (2002) proposed a role for ACC in the appraisal of gains and losses. The contribution of ACC to reward processing makes sense in view of its anatomical position in a network composed by ventral striatum and orbitofrontal cortex and the dense dopaminergic afferences from the ventral tegmental area. As a related matter, studies on drug addiction show abnormal activation of ACC, in particular during state of drug craving (Kilts et al., 2001).

Functional Heterogeneity

Several meta-analyses reveal the heterogeneity of human ACC (Bush et al., 2000; Paus et al., 1998; Peyron et al., 2000; Picard & Strick, 1996). Two major subdivisions have been defined: the most rostral—subgenual—subdivision related to affect, and the caudo-dorsal subdivision related to cognitive and motor functions (Bush et al., 2000). The later has been further divided into three small areas as equivalents of monkey CMAs, with the most anterior one, homologous to CMAr, corresponding to the cognitive ACC (Picard & Strick, 1996). The heterogeneity in humans appears very coherent with the maps described in monkeys. However, the functions supported by each subdivisions and the putative link between them is still a matter of debate. Paus (2001) proposes that the heterogeneity reflects the overlap of three domains (motor control, cognition, and arousal/drive state), placing the ACC in a unique position to translate intentions to actions.

In reality, the simplistic rostro-caudal emotion-cognition-motor gradient appears to contain various eye, arm, and verbal motor representations overlapping with autonomic-, attention-, pain-, and affect-related subdivisions (Paus, 2001; Peyron et al., 2000; Picard & Strick, 1996). Moreover, the cingulate cortex can be divided into dorsal and ventral (gyrus) regions that might have distinct functions (Bush et al., 2000; Peyron et al., 2000) (figure 16.4, plate 19).

Thus, remaining major questions are whether the ACC is organized according to precise functional principles, whether subdivisions implement a common unitary function applied to different domains, or whether they participate in separate processes (see conclusion).

Models and Theories of ACC Function

At present the conflict monitoring, the reinforcement learning, and the selection for action hypotheses are among the most acknowledged theories of ACC functions. We will focus here on three connectionist models that describe networks devoted to the adaptation of behavior and in which ACC has a central role (Botvinick et al., 2001; Dehaene, Kerszberg, & Changeux, 1998; Holroyd & Coles, 2002).

ACC is part of a system able to generate automatic responses, and ready to suppress these responses to generate new ones through cognitive control. Norman and Shallice originally described this system using a two-stage model composed of an automatic selector module (the contention scheduling) and a control system (the supervisory attentional system) (Shallice, 1988). While this model and others try to describe what is control and how it is performed, very few deal with how the system is engaged in the control function. In other words, how is the controller controlled? Botvinick and colleagues have proposed a solution to this problem while retaining the

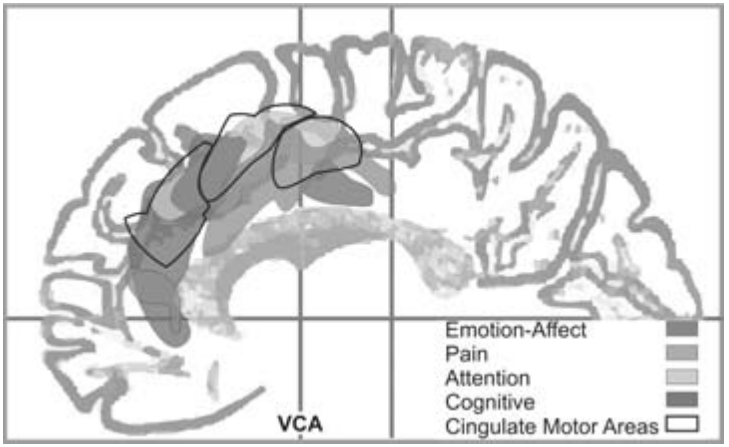


Figure 16.4
Schematic summary of different anterior cingulate zones drawn from results of different meta-analyses. Colored areas regroup zones in which peaks of activations related to different domains have been observed in the literature (see Peyron, Laurent, & Garcia-Larrea, 2000, Swick & Turken, 2002). The three cingulate motor areas defined by Picard & Strick (1996) are represented surrounded by dark lines. See plate 19 for color version.

main features of the Norman-Shallice theory (Botvinick et al., 2001). Here, the lateral prefrontal cortex performs the cognitive control (the maintenance of an adequate representation and inhibition of concurrent representations). Following recurrent correct responses, cognitive control diminishes and automatic responses are more likely to occur. In complex situations, this supposedly leads to conflict between competing response representations, then to behavioral errors, and finally to a need to reengage cognitive control. Botvinick and colleagues propose a conflict-monitoring hypothesis in which a system detects conflict and therefore the need for control, and sends a signal that ultimately will lead to increase in cognitive control. Using data from fMRI and ERP experiments, the authors hypothesize that a subdivision of dorsal ACC is the conflict detector. Indeed, a focalized ACC activation appears in situation where conflict is high (like the incongruent trials in the Stroop test) (Carter et al., 1998; MacDonald, Cohen, Stenger, & Carter, 2000). To account for the ERN, the authors propose that the negative potential is not related to response accuracy itself but to a postresponse conflict between an activated representation of the erroneous response already executed and the coactivated representation of the correct response. Although the model is attractive and explains much of brain imaging data, its main flaw concerns error and reward. It is, for instance, difficult to explain how conflict could occur after presentation of an external error feedback. Moreover, some imaging data suggest

that there is an anatomical dissociation between activation for response competition and error processing (Ullsperger & von Cramon, 2001).

Dehaene and colleagues proposed a model based on two computational spaces. The first, the processing network, encloses functionally specialized subsystems ranging from primary sensory to heteromodal, or high-order categorical, processors (Dehaene, Kerszberg, & Changeux, 1998). The second computational space, the global workspace, consists of distributed neuronal populations connected to the processing network through long-range connections. Interestingly, the global workspace is not a group of fixed cardinal cortical areas but encompass various sets of neurons, in different areas, that are engaged, or configured, at a particular time in a given task. Processors are dynamically engaged by these particular subsets or workspace neurons. Among the five defined specialized processors, the evaluation circuit allows maintenance or change in workspace activity depending on its positive or negative value. The evaluation circuit theoretically includes the ACC. The authors tested a minimal simulation of this architecture on the Stroop task at different stages of training. They show that the global workspace was fully activated during the initial search or learning period, then changed its mode of activation and refined the specificity of its top-down influence on processors during an effortful but efficient execution of the task. Moreover, routinization led to disengagement of the global workspace. Another interesting feature of the model is that repetitive correct performance during effortful execution leads to reduced vigilance and workspace activity and ends up with behavioral errors. The errors in turn induce reactivation of both the vigilance system and workspace activity. This is comparable to the reactivation of cognitive control in the Botvinick et al. model (see above). It is noticeable that the differential involvement of the workspace is highly comparable to cortical, in particular the lateral prefrontal and the ACC, and subcortical activity change in nonroutine and routine situations (Dehaene, Kerszberg, & Changeux, 1998; Jueptner et al., 1997a; Procyk et al., 2000). Moreover, the fact that different subsets of workspace neurons are activated depending on the peripheral processors that must be controlled is consistent with the heterogeneity of the ACC. Although the performance of the model and the properties of workspace neurons are impressive, a few interesting questions remains. Do parts of the workspace endorse specific functions comparable to what is observed in monkey ACC? How does the evaluative circuit relate to the role of dopaminergic neurons in prediction error and to the role of ACC in error processing?

The function of error-related brain activity in the adaptation of behavior might be to signal the occurrence of inefficient responses, and trigger the onset of adaptive strategies through cognitive control and action selection. ACC seems to be a central structure for error processing: it might be the origin of the error-related negativity (ERN), and in monkeys, some ACC unit activities are error-related. Holroyd and Coles propose that the relationships between ACC and the mesocortical dopaminergic

pathway are the source of the ERN (Holroyd & Coles, 2002). In their model, a negative prediction error-signal (inhibition of dopaminergic neurons), which indicates a negative deviation to an expected reward, disinhibits ACC motor neurons, which then produce the cortical error signal. For correct trials, activation of dopaminergic neurons inhibits ACC motor neurons and thus no error signal is produced. This is the reinforcement-learning theory. In this model, the ACC in itself does not compute the error: "In this view, the anterior cingulate cortex acts as a motor control filter, enabling any one of the motor controllers to take command of the motor system" (Holroyd & Coles, 2002). Thus, the ACC will use the error input to shift responsibility for the task to the most adapted processing module. Here the ACC is conceived as a selector. As stated in their paper, Holroyd and Coles do not exclude activation of ACC during conflict detection but challenge the idea that ERN is the manifestation of a conflict. Thus, the reinforcement-learning theory is specific to error processing and does not really explain ACC activity in conflicting situations.

Summary and Conclusions

Although human and nonhuman primate data show strong similarities, one has to be careful when comparing functional imaging and local electrophysiological results. fMRi and ERP measure global brain structure activation without giving clues about what information neurons are processing at small temporal and anatomical scales. For instance, activation of one neuronal population after error and activation of another population (but from the same area) after reward might lead to identical global signals (Bush et al., 2002). Looking at unit activity may lead to propose a role for the area in two potentially related functions, whereas looking at global signals may lead to the construction of one intermediate function. The best approach might be to combine both techniques in nonhuman primates and thereby fill the gap between human brain imaging and monkey unit recordings.

There is a consensus that the ACC possesses an evaluative, or performance monitoring, function either for conflict, error, or reward processing. Although various ACC activations are observed in different contexts, one may consider that these results reveal a functional principle. Botvinick and colleagues note, for instance, that error, pain, and conflict are all devoted to signaling that something has gone or will go wrong, and can induce a shift in attentional or cognitive resources to react and adapt behaviour (Botvinick et al., 2001). Existing models suggest that ACC participates in a comparison process, indirectly through the detection of coactivations (Botvinick et al., 2001), or directly by comparing representations (e.g., produced response vs. representation of correct response, actual consequences of action vs. expected consequences).

It is often difficult to directly relate ACC neurophysiological data in monkeys with theories from the brain-imaging literature. One example concerns the conflict-monitoring hypothesis. Although activity related to free selection or to trial-and-error learning (Procyk et al., 2000; Shima et al., 1991) could reflect conflict detection, so far ACC studies in monkeys did not use tasks strictly comparable to the human Stroop test. In monkey studies, data show that dorsal anterior cingulate neurons as a population (mostly in and anterior to CMAr), encode for expected reward or goal, and for errors. This is also observed in human studies (Bachevalier, Meunier, Lu, & Ungerleider, 1997; Bush et al., 2002; Ullsperger & von Cramon, 2003). As underlined by Gehring and Willoughby, these data support a role of ACC in the appraisal of outcomes (Gehring & Willoughby, 2002). In fact, cingulate unit activity participates in the representation of external and internal events (stimulus onset, movement preparation) but mainly in relation to their expected or actual value, for example, in terms of distance to or quality of the incoming reward. Possibly, the integration of goal (expected positive outcome) representation to events and motor representations, not only gives a behavioral value to an event but also computationally prepares the structure to process deviation of the actual from the expected outcome. Here again the ACC participates in a comparison process that enables the brain to compare outcomes to predictions and, depending on the result, choose to stay or shift behavioral responses. In this view however, and in opposition to models described above, the ACC has itself the means to compute these comparisons.

To conclude, we should keep in mind that ACC functional specificities are related to its position in particular structured networks. For instance, the strength of relationships between ACC and the lateral prefrontal cortex is one remarkable feature of intrafrontal connectivity (Paus, 2001), and is central to the integrity of cognitive processes devoted to the adaptation of behavior. Further along the line, the study of the interdependence between these structures and aminergic systems, obviously devoted to learning processes and cognitive functions, will be an essential direction for future research devoted to the neurobiological bases of executive functions.

References

Arikuni, T., & Kubota, K. (1986). The organization of prefrontocaudate projections and their laminar origin in the macaque monkey: a retrograde study using HRP-gel. *Journal of Comparative Neurology*, *244*(4), 492–510.

Arikuni, T., Sakai, M., & Kubota, K. (1983). Columnar aggregation of prefrontal and anterior cingulate cortical cells projecting to the thalamic mediodorsal nucleus in the monkey. *Journal of Comparative Neurology*, *220*(1), 116–125.

Bachevalier, J., Meunier, M., Lu, M., & Ungerleider, L. G. (1997). Thalamic and temporal cortex input to medial prefrontal cortex in rhesus monkeys. *Experimental Brain Research*, *115*(3), 430–444.

Barbas, H., & Mesulam, M. M. (1985). Cortical afferent input to the principalis region of the rhesus monkey. *Neuroscience*, *15*(3), 619–637.

Barbas, H., & Pandya, D. N. (1989). Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. *Journal of Comparative Neurology*, *286*(3), 353–375.

Bates, J. F., & Goldman-Rakic, P. S. (1993). Prefrontal connections of medial motor areas in the rhesus monkey. *Journal of Comparative Neurology*, *336*(2), 211–228.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1–3), 7–15.

Bench, C. J., Frith, C. D., Grasby, P. M., Friston, K. J., Paulesu, E., Frackowiak, R. S., et al. (1993). Investigations of the functional anatomy of attention using the stroop test. *Neuropsychologia*, *31*(9), 907–922.

Blakemore, S. J., Rees, G., & Frith, C. D. (1998). How do we predict the consequences of our actions? A functional imaging study. *Neuropsychologia*, *36*(6), 521–529.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652.

Brooks, V. B. (1986). How does the limbic system assist motor learning? A limbic comparator hypothesis. *Brain, Behavior & Evolution*, *29*, 29–53.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trendsin Cognitive Science*, *4*(6), 215–222.

Bush, G., Vogt, B. A., Holmes, J., Dale, A. M., Greve, D., Jenike, M. A., et al. (2002). Dorsal anterior cingulate cortex: a role in reward-based decision making. *Proceedings of the National Academy of Sciences*, *99*(1), 523–528.

Cadoret, G., & Smith, A. M. (1997). Comparison of the neuronal activity in the SMA and the ventral cingulate cortex during prehension in the monkey. *Journal of Neurophysiology*, *77*(1), 153–166.

Carmichael, S. T., & Price, J. L. (1994). Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *Journal of Comparative Neurology*, *346*(3), 366–402.

Carmichael, S. T., & Price, J. L. (1995). Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *Journal of Comparative Neurology*, *363*, 642–664.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, *280*(5364), 747–749.

Chen, Y. C., Thaler, D., Nixon, P. D., Stern, C. E., & Passingham, R. E. (1995). The functions of the medial premotor cortex. II. The timing and selection of learned movements. *Experimental Brain Research*, *102*(3), 461–473.

Cipolloni, P. B., & Pandya, D. N. (1999). Cortical connections of the frontoparietal opercular areas in the rhesus monkey. *Journal of Comparative Neurology*, *403*(4), 431–458.

Cohen, R. A., Kaplan, R. F., Moser, D. J., Jenkins, M. A., & Wilkinson, H. (1999a). Impairments of attention after cingulotomy. *Neurology*, *53*(4), 819–824.

Cohen, R. A., Kaplan, R. F., Zuffante, P., Moser, D. J., Jenkins, M. A., Salloway, S., et al. (1999b). Alteration of intention and self-initiated action associated with bilateral anterior cingulotomy. *Journal of Neuropsychiatry and Clinical Neuroscience*, *11*(4), 444–453.

Crino, P. B., Morrison, J. H., & Hof, P. R. (1993). Monoaminergic innervation of cingulate cortex. In B. A. Vogt & M. Gabriel (Eds.), *Neurobiology of Cingulate Cortex and Limbic Thalamus* (pp. 285–310). Basel: Birkaüser.

Dehaene, S., Kerszberg, M., & Changeux, J. P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences*, *95*(24), 14529–14534.

Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science*, *5*(5), 303–305.

Deiber, M. P., Passingham, R. E., Colebatch, J. G., Friston, K. J., Nixon, P. D., & Frackowiak, R. S. (1991). Cortical areas and the selection of movement: A study with positron emission tomography. *Experimental Brain Research*, *84*, 393–402.

Devinksy, O., & Luciano, D. (1993). The contributions of cingulate cortex to human behavior. In B. A. Vogt & M. Gabriel (Eds.), *Neurobiology of Cingulate Cortex and Limbic Thalamus* (pp. 528–556). Basel: Birkhäuser.

Dum, R. P., & Strick, P. L. (1991). The origin of corticospinal projections from the premotor areas in the frontal lobe. *Journal of Neuroscience*, *11*(3), 667–689.

Dum, R. P., & Strick, P. L. (1993). Cingulate motor areas. In B. A. Vogt & M. Gabriel (Eds.), *Neurobiology of Cingulate Cortex and Limbic Thalamus* (pp. 415–441). Basel: Birkhaüser.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trendsin Neuroscience*, *23*(10), 475–483.

Frankenstein, U. N., Richter, W., McIntyre, M. C., & Remy, F. (2001). Distraction modulates anterior cingulate gyrus activations during the cold pressor test. *NeuroImage*, *14*(4), 827–836.

Frith, C. D., Friston, K. J., Liddle, P. F., & Frackowiak, R. S. (1991). Willed action and the prefrontal cortex in man: A study with PET. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, *244*, 241–246.

Funahashi, S. (2001). Neuronal mechanisms of executive control by the prefrontal cortex. *Neuroscience Research*, *39*(2), 147–165.

Gaymard, B., Rivaud, S., Cassarini, J. F., Dubard, T., Rancurel, G., Agid, Y., et al. (1998). Effects of anterior cingulate cortex lesions on ocular saccades in humans. *Experimental Brain Research*, *120*(2), 173–183.

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, *295*(5563), 2279–2282.

George, M. S., Ketter, T. A., Parekh, P. I., Rosinsky, N., Ring, H., Casey, B. J., et al. (1994). Regional brain activity when selecting a response despite interference: an $H_2{}^{15}O$ PET study of the Stroop and an emotional stroop. *Human Brain Mapping*, *1*, 194–209.

Grafton, S. T., Hazeltine, E., & Ivry, R. B. (1995). Functional mapping of sequence learning in normal humans. *Journal Cognitive Neuroscience*, *7*(4), 497–510.

Hadland, K. A., Rushworth, M. F., Gaffan, D., & Passingham, R. E. (2003). The anterior cingulate and reward-guided selection of actions. *Journal of Neurophysiology*, *89*(2), 1161–1164.

He, S. Q., Dum, R. P., & Strick, P. L. (1995). Topographic organization of corticospinal projections from the frontal lobe: motor areas on the medial surface of the hemisphere. *Journal of Neuroscience*, *15*(5 Pt 1), 3284–3306.

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709.

Huerta, M. F., Krubitzer, L. A., & Kaas, J. H. (1987). Frontal eye field as defined by intracortical microstimulation in squirrel monkeys, owl monkeys, and macaque monkeys. II. Cortical connections. *Journal of Comparative Neurology*, *265*(3), 332–361.

Insausti, R., Amaral, D. G., & Cowan, W. M. (1987). The entorhinal cortex of the monkey: II. Cortical afferents. *Journal of Comparative Neurology*, *264*(3), 356–395.

Ito, S., Stuphorn, V., Brown, J. W., & Schall, J. D. (2003). Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science*, *302*(5642), 120–122.

Jenkins, I. H., Brooks, D. J., Nixon, P. D., Frackowiak, R. S., & Passingham, R. E. (1994). Motor sequence learning: a study with positron emission tomography. *Journal of Neuroscience*, *14*(6), 3775–3790.

Jueptner, M., Frith, C. D., Brooks, D. J., Frackowiak, R. S., & Passingham, R. E. (1997a). Anatomy of motor learning. II. Subcortical structures and learning by trial and error. *Journal of Neurophysiology*, *77*(3), 1325–1337.

Jueptner, M., Stephan, K. M., Frith, C. D., Brooks, D. J., Frackowiak, R. S., & Passingham, R. E. (1997b). Anatomy of motor learning. I. Frontal cortex and attention to action. *Journal of Neurophysiology*, *77*(3), 1313–1324.

Kilts, C. D., Schweitzer, J. B., Quinn, C. K., Gross, R. E., Faber, T. L., Muhammad, F., et al. (2001). Neural activity related to drug craving in cocaine addiction. *Archives of General Psychiatry*, *58*(4), 334–341.

Kunishio, K., & Haber, S. N. (1994). Primate cingulostriatal projection: Limbic striatal versus sensorimotor striatal input. *Journal of Comparative Neurology*, *350*(3), 337–356.

Lu, M. T., Preston, J. B., & Strick, P. L. (1994). Interconnections between the prefrontal cortex and the premotor areas in the frontal lobe. *Journal of Comparative Neurology*, *341*, 375–392.

Luppino, G., Matelli, M., Camarda, R., & Rizzolatti, G. (1993). Corticocortical connections of area F3 (SMA-proper) and area F6 (pre-SMA) in the macaque monkey. *Journal of Comparative Neurology*, *338*(1), 114–140.

Luppino, G., Matelli, M., Camarda, R., & Rizzolatti, G. (1994). Corticospinal projections from mesial frontal and cingulate areas in the monkey. *NeuroReport*, *5*(18), 2545–2548.

Luppino, G., Matelli, M., Camarda, R. M., Gallese, V., & Rizzolatti, G. (1991). Multiple representations of body movements in mesial area 6 and the adjacent cingulate cortex: An intracortical microstimulation study in the macaque monkey. *Journal of Comparative Neurology*, *311*(4), 463–482.

Luppino, G., Rozzi, S., Calzavara, R., & Matelli, M. (2003). Prefrontal and agranular cingulate projections to the dorsal premotor areas F2 and F7 in the macaque monkey. *European Journal of Neuroscience*, *17*(3), 559–578.

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, *288*(5472), 1835–1838.

Marconi, B., Genovesio, A., Battaglia-Mayer, A., Ferraina, S., Squatrito, S., Molinari, M., et al. (2001). Eye-hand coordination during reaching. I. Anatomical relationships between parietal and frontal cortex. *Cerebral Cortex*, *11*(6), 513–527.

Matelli, M., Luppino, G., & Rizzolatti, G. (1991). Architecture of superior and mesial area 6 and the adjacent cingulate cortex in the macaque monkey. *Journal of Comparative Neurology*, *311*(4), 445–462.

McGuire, P. K., Bates, J. F., & Goldman-Rakic, P. S. (1991a). Interhemispheric integration: I. Symmetry and convergence of the corticocortical connections of the left and the right principal sulcus (PS) and the left and the right supplementary motor area (SMA) in the rhesus monkey. *Cerebral Cortex*, *1*(5), 390–407.

McGuire, P. K., Bates, J. F., & Goldman-Rakic, P. S. (1991b). Interhemispheric integration: II. Symmetry and convergence of the corticostriatal projections of the left and the right principal sulcus (PS) and the left and the right supplementary motor area (SMA) of the rhesus monkey. *Cerebral Cortex*, *1*(5), 408–417.

Morecraft, R. J., Geula, C., & Mesulam, M. M. (1993). Architecture of connectivity within a cingulo-fronto-parietal neurocognitive network for directed attention. *Archives of Neurology*, *50*(3), 279–284.

Morecraft, R. J., & Van Hoesen, G. W. (1992). Cingulate input to the primary and supplementary motor cortices in the rhesus monkey: evidence for somatotopy in areas 24c and 23c. *Journal of Comparative Neurology*, *322*(4), 471–489.

Morecraft, R. J., & Van Hoesen, G. W. (1993). Frontal granular cortex input to the cingulate (M3), supplementary (M2) and primary (M1) motor cortices in the rhesus monkey. *Journal of Comparative Neurology*, *337*(4), 669–689.

Morecraft, R. J., & Van Hoesen, G. W. (1998). Convergence of limbic input to the cingulate motor cortex in the rhesus monkey. *Brain Research Bulletin*, *45*(2), 209–232.

Niki, H., & Watanabe, M. (1976). Cingulate unit activity and delayed response. *Brain Research*, *110*, 381–386.

Pandya, D. N., Dye, P., & Butters, N. (1971). Efferent cortico-cortical projections of the prefrontal cortex in the rhesus monkey. *Brain Research*, *31*(1), 35–46.

Pandya, D. N., & Vignolo, L. A. (1971). Intra- and interhemispheric projections of the precentral, premotor and arcuate areas in the rhesus monkey. *Brain Research*, *26*(2), 217–233.

Pardo, J. V., Pardo, P. J., Janer, K. W., & Raichle, M. E. (1990). The anterior cingulate cortex mediates processing selection in the stroop attentional conflict paradigm. *Proceedings of the National Academy of Sciences*, *87*, 256–259.

Paus, T. (2001). Primate anterior cingulate cortex: Where motor control, drive and cognition interface. *Nature Neuroscience*, *2*(6), 417–424.

Paus, T., Koski, L., Caramanos, Z., & Westbury, C. (1998). Regional differences in the effects of task difficulty and motor output on blood flow response in the human anterior cingulate cortex: A review of 107 PET activation studies. *NeuroReport*, *9*(9), R37–47.

Paus, T., Petrides, M., Evans, A. C., & Meyer, E. (1993). Role of the human anterior cingulate cortex in the control of oculomotor, manual, and speech responses: A positron emission tomography study. *Journal of Neurophysiology*, *70*(2), 453–469.

Petit, L., Courtney, S. M., Ungerleider, L. G., & Haxby, J. V. (1998). Sustained activity in the medial wall during working memory delays. *Journal of Neuroscience*, *18*(22), 9429–9437.

Peyron, R., Laurent, B., & Garcia-Larrea, L. (2000). Functional imaging of brain responses to pain. A review and meta-analysis (2000). *Neurophysiology Clinical*, *30*(5), 263–288.

Phelps, E. A., Hyder, F., Blamire, A. M., & Shulman, R. G. (1997). FMRI of the prefrontal cortex during overt verbal fluency. *NeuroReport*, *8*(2), 561–565.

Picard, N., & Strick, P. L. (1996). Motor areas of the medial wall: A review of their location and functional activation. *Cerebral Cortex*, *6*(3), 342–353.

Picard, N., & Strick, P. L. (1997). Activation on the medial wall during remembered sequences of reaching movements in monkeys. *Journal of Neurophysiology*, *77*(4), 2197–2201.

Posner, M. I., & DiGirolamo, G. J. (1998). Executive attention: Conflict, target detection and cognitive control. In R. Parasuraman (Ed.), *The Attentive Brain* (pp. 401–424). Cambridge, MA: MIT Press.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42.

Procyk, E., & Joseph, J. P. (2001). Characterization of serial order encoding in the monkey anterior cingulate sulcus. *European Journal of Neuroscience*, *14*(6), 1041–1046.

Procyk, E., Tanaka, Y. L., & Joseph, J. P. (2000). Anterior cingulate activity during routine and non-routine sequential behaviors in macaques. *Nature Neuroscience*, *3*(5), 502–508.

Raichle, M. E., Fiez, J. A., Videen, T. O., MacLeod, A. K., Pardo, J. V., Fox, P. T., et al. (1994). Practice-related changes in human brain functional anatomy during non-motor learning. *Cerebral Cortex*, *4*, 8–26.

Rosene, D. L., & Van Hoesen, G. W. (1977). Hippocampal efferents reach widespread areas of cerebral cortex and amygdala in the rhesus monkey. *Science*, *198*(4314), 315–317.

Rushworth, M. F., Hadland, K. A., Gaffan, D., & Passingham, R. E. (2003). The effect of cingulate cortex lesions on task switching and working memory. *Journal of Cognitive Neuroscience*, *15*(3), 338–353.

Schall, J. D., Stuphorn, V., & Brown, J. W. (2002). Monitoring and control of action by the frontal lobes. *Neuron*, *36*(2), 309–322.

Selemon, L. D., & Goldman-Rakic, P. S. (1988). Common cortical and subcortical targets of the dorsolateral prefrontal and posterior parietal cortices in the rhesus monkey: Evidence for a distributed neural network subserving spatially guided behavior. *Journal of Neuroscience*, *8*(11), 4049–4068.

Shallice, T. (1988). *From Neuropsychology to Mental Structure*. New York: Cambridge University Press.

Shidara, M., & Richmond, B. J. (2002). Anterior cingulate: Single neuronal signals related to degree of reward expectancy. *Science*, *296*(5573), 1709–1711.

Shima, K., Aya, K., Mushiake, H., Inase, M., Aizawa, H., & Tanji, J. (1991). Two movement-related foci in the primate cingulate cortex observed in signal-triggered and self-paced forelimb movements. *Journal of Neurophysiology*, *65*(2), 188–202.

Shima, K., & Tanji, J. (1998). Role for cingulate motor area cells in voluntary movement selection based on reward. *Science*, *282*(5392), 1335–1338.

Siwek, D. F., & Pandya, D. N. (1991). Prefrontal projections to the mediodorsal nucleus of the thalamus in the rhesus monkey. *Journal of Comparative Neurology*, *312*(4), 509–524.

Stern, C. E., & Passingham, R. E. (1994). The nucleus accumbens in monkeys (*Macaca fascicularis*): I. The organization of behaviour. *Behavioral Brain Research*, *61*(1), 9–21.

Swick, D., & Turken, A. U. (2002). Dissociation between conflict detection and error monitoring in the human anterior cingulate cortex. *Proceedings of the National Academy of Sciences*, *99*(25), 16354–16359.

Thaler, D., Chen, Y. C., Nixon, P. D., Stern, C. E., & Passingham, R. E. (1995). The functions of the medial premotor cortex. I. Simple learned movements. *Experimental Brain Research*, *102*(3), 445–460.

Turken, A. U., & Swick, D. (1999). Response selection in the human anterior cingulate cortex. *Nature Neuroscience*, *2*(10), 920–924.

Ullsperger, M., & von Cramon, D. Y. (2001). Subprocesses of performance monitoring: A dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *NeuroImage*, *14*(6), 1387–1401.

Ullsperger, M., & von Cramon, D. Y. (2003). Error monitoring using external feedback: Specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. *Journal of Neuroscience*, *23*(10), 4308–4314.

Vidal, F., Hasbroucq, T., Grapperon, J., & Bonnet, M. (2000). Is the "error negativity" specific to errors? *Biological Psychology*, *51*(2–3), 109–128.

Vogt, B. A. (1993). Structural organization of cingulate cortex: areas, neurons, and somatodendritic transmitter receptors. In B. A. Vogt & M. Gabriel (Eds.), *Neurobiology of Cingulate Cortex and Limbic Thalamus* (pp. 19–70). Basel: Birkhaüser.

Vogt, B. A., Finch, D. M., & Olson, C. R. (1992). Functional heterogeneity in cingulate cortex: The anterior executive and posterior evaluative regions. *Cerebral Cortex*, *2*, 435–443.

Vogt, B. A., Nimchinsky, E. A., Vogt, L. J., & Hof, P. R. (1995). Human cingulate cortex: Surface features, flat maps, and cytoarchitecture. *Journal of Comparative Neurology*, *359*(3), 490–506.

Vogt, B. A., Pandya, D. N., & Rosene, D. L. (1987). Cingulate cortex of the rhesus monkey: I. Cytoarchitecture and thalamic afferents. *Journal of Comparative Neurology*, *262*(2), 256–270.

Wang, Y., Shima, K., Sawamura, H., & Tanji, J. (2001). Spatial distribution of cingulate cells projecting to the primary, supplementary, and pre-supplementary motor areas: A retrograde multiple labeling study in the macaque monkey. *Neuroscience Research*, *39*(1), 39–49.

V Visual Representations and the Temporal Lobe

17 Does the Human Brain Process Objects of Expertise Like Faces? A Review of the Evidence

Elinor McKone and Nancy Kanwisher

In the quest for homologies between monkey and human cognition, domain-specific mechanisms are excellent candidates: highly specialized processors that operate on specific kinds of information, that develop early, and that are likely to be evolutionarily conserved. Here we explore one of the strongest contenders for a domain-specific processor, the human system for the perceptual analysis of faces.

Substantial evidence supports the domain specificity of face processing in humans. First, behavioral data indicate a different style of cognitive processing for faces (configural or holistic) than for other objects (feature- or part-based); for example, effects of orientation inversion are much more severe for faces than for other objects (e.g., Yin, 1969), and parts appear to be particularly strongly integrated into wholes in upright faces (e.g., Tanaka & Farah, 1993). Second, brain injury can produce selective deficits in face recognition while leaving object discrimination intact (McNeil & Warrington, 1993; Sergent & Signoret, 1992; de Renzi, 1986; Wada & Yamamoto, 2001) and vice versa (Moscovitch, Winocur, & Behrmann, 1997). Third, brain imaging (Kanwisher, McDermott, & Chun, 1997; McCarthy, Puce, Gore, & Allison, 1997) and event-related potential (ERP) research (Bentin et al., 1996) have found neural responses that are highly specific to faces (see also Desimone, Albrignt, Gross, & Bruce, 1984). These findings support the "face specificity" hypothesis, according to which the cognitive and neural mechanisms underlying face processing are selectively engaged in perceptually processing faces per se and play little if any role in the perceptual analysis of nonface stimuli.

However, the face specificity hypothesis has become the subject of considerable debate. In the alternative "expertise" hypothesis, it is argued that mechanisms for face processing are not engaged only by faces, but are also applied in expert within-class discrimination of nonfaces (Diamond & Carey, 1986; Gauthier & Tarr, 1997). The idea here is that faces all share a basic "first-order" configuration, that is, the same parts are arranged in the same fixed layout (e.g., two eyes above a central nose), but recognition requires discrimination of individuals (*Bill* vs. *John*) who differ only in "second-order" deviations from this shared basic structure. In contrast, in most object

recognition situations only between-class discrimination is required (*dog* vs. *bird*), for which parts and their first-order configurations are sufficient for identification. According to the expertise hypothesis, then, face recognition seems special only because adult humans have had substantial experience discriminating individual faces, but almost no practice at making similar within-class discriminations about objects. This view predicts that, for those rare people who have developed expertise in within-class object discrimination—such as dog-show judges capable of discriminating one Scotch terrier from another (Diamond and Carey, 1986)—the same mechanisms will be engaged in the processing of faces and of objects-of-expertise.

The debate between the face specificity and expertise hypotheses has now been active for some years. The aim of the present chapter is to provide a critical review of the current empirical evidence on the issue. We focus here on the two key predictions of the expertise hypothesis that have received most attention: that experts with nonface objects (1) should show behavioral signatures of face processing when processing nonface objects of expertise, and (2) should engage putatively face-specific neural mechanisms when perceptually processing objects of expertise. As our review of the literature shows, there is currently little convincing evidence for these predictions, and considerable evidence against them. Thus current evidence favors the face specificity hypothesis, although several key experiments remain to be conducted. We do not address here the important but distinct question of the developmental origins of face processing mechanisms, that is, the relative roles of genetic and experiential factors in their development.

Behavioral Evidence

How Faces Are Special in Human Behavior: Configural Processing

In behavioral terms, the special style of cognitive processing that occurs for faces is variously referred to as configural, holistic, or relational processing. There is little agreement on the exact nature of this type of processing (Maurer, Le Grand, & Mondloch, 2002), but the general idea is that either there are interactions between multiple parts over a broad region of the face ("configural," cf. Rhodes, 1988), or there is no decomposition into parts at all ("holistic"; Tanaka & Sengco, 1997), beyond perhaps simple lines and edges of early vision (Moscovitch, Winocur, & Behrmann, 1997). It is also presumed that configural processing includes coding of detailed distance information, such as the spacing between the eyes (e.g., "relational," Diamond & Carey, 1986). For present purposes, we employ the term *configural* and do not distinguish between alternative definitions. Configural processing is usually contrasted with "part-based" processing, which is presumed to involve some sort of decomposition at obvious part boundaries, plus some coding of basic relations among these parts (above, below, to-the-side-of, etc).

Initial behavioral evidence for special processing of faces came from the *disproportionate inversion effects* in recognition memory. All objects are remembered worse inverted than upright but inversion effects are much larger for faces (25 percentage point decrement) than for a wide range of other object classes (2–10 points)(Yin, 1969; Diamond & Carey, 1986; Scapinello & Yarmey, 1970). This finding occurs despite the memory-task requirement for within-class discrimination in all cases.

Other methods directly demonstrate configural processing of upright faces. In the *composite effect* (Young, Hellawell, & Hay, 1987), it is difficult to name one half of a face formed from a composite of two different individuals when the two halves are aligned (indicating automatic perceptual integration into a new whole), in comparison to a control condition in which the two halves are offset (figure 17.1). In the *part-whole effect* (Tanaka & Farah, 1993), memory for a face part (Bill's nose) is much poorer in isolation (Bill's nose vs. John's nose) than in the context of the original whole face (Bill's nose in Bill's face vs. John's nose in Bill's face). In the *whole-vs-configurally-transformed-whole* version of the part-whole paradigm (Tanaka & Sengco, 1997), memory for a part is worse in the context of a configurally altered version of the original face (Bill's face with the eyes shifted apart slightly) than in the original face. In other *relational alteration paradigms*, perceived bizarreness, perceived distinctiveness, and memory are substantially affected by altering distances between natural face parts (shifting eyes apart, or the mouth down; Bartlett & Searcy, 1993; Leder & Bruce, 1998; Le Grand, Mondloch, Maurer, & Brent, 2001; Rhodes, Brake, & Atkinson, 1993).

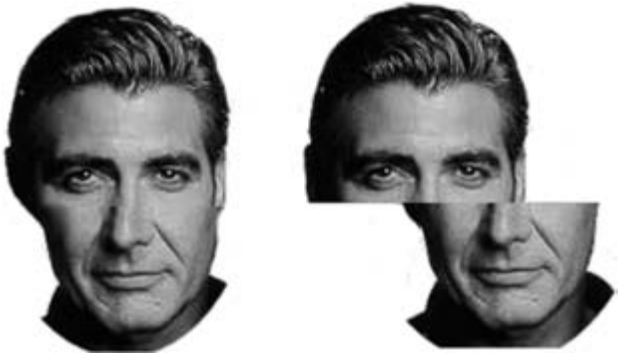


Figure 17.1
The composite effect of Young et al. (1987). To name the top half as George Clooney (or the bottom half as Harrison Ford) takes approximately 200 ms longer in the aligned version on the left than in the misaligned version on the right. Given the two conditions are matched for simple response competition (the to-be-ignored half is always a different individual), this effect must reflect perceptual integration of the two halves into a new whole in the aligned condition. The composite effect does not occur when the stimuli are shown inverted.

In all the studies cited above, configural effects occurred for upright faces but were absent or nearly absent for inverted faces. To the extent that the paradigms have been applied to objects, nonexperts have also demonstrated little or no indication of configural processing (houses, dogs, car fronts, and biological cells) in the part-whole paradigm (Tanaka & Sengco, 1997, Tanaka et al., 1996, cited in Tanaka and Gauthier 1997).

Behavioral Findings: Laboratory-Trained Experts

One approach to expertise has been to test laboratory-trained subjects, using an artificial class of objects called "greebles" (figure 17.2). Participants are trained to identify many greebles over several sessions (8–10 hr) until reaction times (RTs) for individual identity decisions are as fast as those for "family" and "gender" decisions. We have general theoretical concerns about experiments involving greebles. Any *failure* to find evidence of facelike special processing with greebles is not conclusive, given the relatively small amounts of practice and the rather weak criterion for expertise (cf. 31 years to show face-sized inversion effects in behavioral studies; see Diamond & Carey, 1986, below). Equally, however, any *positive evidence* of "face-specific" processing for greebles can be inconclusive. This is because greebles have a high degree of structural similarity to faces and/or face-body combinations (figure 17.2). Thus small greeble expertise effects could reflect a specialized face-processing mechanism learning to stretch its definition of a face, rather than a generic expertise effect that could occur for any object class. However, because others are taking the results of greeble studies as relevant to the debate, we will consider them in detail here.

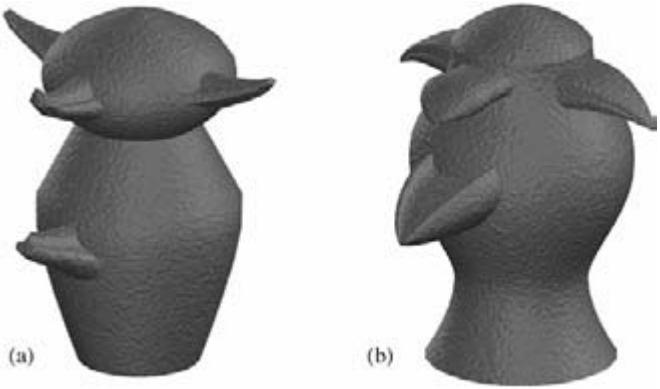


Figure 17.2
Greebles have two horizontally arranged parts above two centrally arranged parts, replicating the T-shaped configuration of eyes, nose and mouth. Perceptually, most appear to have heads (*a*), and in some cases the whole greeble looks like a face (*b*). (Images provided courtesy of Michael J. Tarr, Brown University.)

Several studies have examined behavioral performance in subjects trained with Greebles. The authors of those studies have summarized their findings as having "compared a wide range of putatively face-specific behavioral effects across Greeble novices and Greeble experts and . . . consistently obtained face-like patterns of performance with experts but not novices" (Tarr, 2003), or as "suggest[ing] that subjects shifted from feature-based to more configural processing as they became greeble experts" (Gauthier & Tarr, 1997; Gauthier et al., 1999). Gauthier and Tarr have thus argued strongly for the expertise hypothesis.

Our own review of the greeble papers, however, indicates that the evidence is in fact very weak. The essential problem is that Gauthier and Tarr have focused on a few close-to-significant positive findings, and have ignored equally important results that were either null or in the reverse-to-predicted direction. To explain this, it is necessary to go into some detail and describe the results in full.[1]

In the *standard part-whole paradigm*, Gauthier and Tarr (1997) found a small whole-over-part advantage for both novices (6 percentage points) and experts (11 percentage points).The effect reached significance only for experts but, at the same time, there was no interaction involving expert vs. novice status. In Gauthier, Williams, Tarr, and Tanaka (1998), experts showed a significant whole-over-part advantage for one of three different parts tested, but so did novices. Experts showed trends in the reverse-to-predicted direction for the other two parts. Averaged across all parts, the mean part-whole difference trended in the wrong direction for an expertise effect (7 percentage points for novices, vs. 0 percentage points for experts). In Gauthier and Tarr (2002), the size of the part-whole difference also failed to increase with expertise across five training sessions (d′ difference in session 1 = .76, session 5 = .68).

Comparing *part-in-original-whole* with *part-in-configurally-transformed-whole*, Gauthier and Tarr (1997) found a significant difference for experts on reaction time (although there was no effect at all on accuracy), with a close-to-significant interaction involving expert-vs.-novice status. While these results were suggestive of an expertise effect, further studies failed to confirm the finding. In Gauthier, Williams, Tarr, and Tanaka (1998) only one part showed a close-to-significant effect, while the other two parts showed trends in the reverse-to-predicted direction. Averaged across all parts, the mean effect was 0 percentage points for novices and 0 percentage points for experts (cf., 5–11 points for faces; Tanaka and Sengco, 1997). Gauthier and Tarr (2002) again tested all three parts, on both accuracy and RT. In experts, there was one significant effect in the predicted direction (RT for the quiff), but there was also another significant effect in the reverse-to-predicted direction (accuracy for the dunth). Averaged over all parts, the size of the effect in d′ accuracy was 0.69 in session 1 (novices) and 0.64 in session 5 (experts).

In the *composite test*, Gauthier, Williams, Tarr, and Tanaka (1998) found that, for composites made of same-family halves, there was a nonsignificant trend in the

predicted direction (i.e., aligned more difficult than unaligned) on reaction time in greeble experts, and no effect on accuracy. For composites made of different-family halves, trends were in the reverse-to-predicted direction on both accuracy and reaction time. In Gauthier and Tarr (2002), after five practice sessions, there was no indication that experts had developed a composite effect. On reaction times, the aligned vs. unaligned difference was 12 ms (in the predicted direction, but with SEMs for each condition = 35 ms) and the difference on d′ accuracy was .31 in the reverse-to-predicted direction. Statistics showed a close-to-significant interaction with degree of training on reaction times (there was no effect on accuracy), but this reflected a peculiar finding that the composite effect started close to zero, went in the reverse-to-predicted direction in sessions 2 and 3, and then returned to zero.

In interpreting these findings, Gauthier and Tarr (2002) accept that there is *no part-whole effect* for greeble experts. Rather than take this as evidence against configural processing, however, they choose to define the standard part-whole test as providing a general test of contextual influences rather than a test of configural processing. We see no theoretical basis for this idea. We agree that most objects show some advantage for parts in the context of wholes rather than in isolation (and the same is true for words; i.e., "the word superiority effect" [Reicher, 1969; Wheeler, 1970]), but this ignores the fact that the part-whole effect remains very much greater for faces than it is for objects. Gauthier and Tarr do accept the *whole-vs.-configurally-transformed-whole* test and the *composite paradigm* as tests of configural processing. However, these measures do not show evidence of configural processing in greeble experts, as discussed above.

We also note that most greeble studies have not explicitly tested for *disproportionate inversion effects* in experts. Two experiments including inverted greebles do not suggest any unusually large (i.e., face-sized) inversion effects. In part identification in the context of the whole studied stimulus, greeble experts showed only a 4-percentage-point decrement with inversion (87 percent correct upright, 83 percent inverted [Gauthier & Tarr, 1997]); this compares with an 18-percentage-point decrement previously reported for faces (87 percent upright, 69 percent inverted [Tanaka & Sengco, 1997]). Second, in naming whole greebles that were learned upright, greeble experts showed a misorientation effect (60, 120, and 180 degree image-plane rotations were combined) that trended in the reverse-to-predicted direction for an expertise effect (RT difference between upright and misoriented = 180 ms for experts, 231 ms for novices [Gauthier, Williams, Tarr, & Tanaka, 1998]).[2] In a direct test, Rossion et al. (2002) found a larger inversion effect after training with greebles (46 ms) than before (25 ms), although the effect remained smaller than for faces (75 ms).

Behavioral Findings: Real-World Experts

As mentioned above, lab training studies provide only around 10 hours of experience, which is very different from the lifetime of experience all of us have with faces. In

another approach to investigating expertise which gets a little closer to the situation with faces, several studies have tested real world experts, such as dog show judges with 5, 10, 20, or more years of experience in individuating exemplars of dogs within their particular breed-of-expertise.

In an early indication that configural processing might develop for objects-of-expertise, Diamond and Carey (1986) found that dog show judges (with 31 years experience) showed face-sized inversion effects on memory for their breed of expertise. When a wider range of breeds were used, the inversion effect for dogs remained smaller than that for faces, and the inversion × expertise interaction did not reach significance. These results were taken to indicate that facelike processing of objects (a) was possible and (b) took many years of extensive experience to develop. We note, however, that memory in general is better when appropriate preexisting knowledge can be applied to the to-be-learned material (e.g., soccer experts show better memory for text describing a soccer game than do novices; Schneider, Korkel, & Weinert, 1989). Thus, even though Diamond and Carey's dog experts were clearly applying some sort of perceptual knowledge better to upright dogs than to inverted dogs, it is not necessarily the case that this knowledge was configural in nature. However, a recent study using a perceptual task (sequential matching) found face-sized inversion effects for discrimination of body configuration (Reed et al., 2003), which could reflect either widespread expertise for bodies, or the engagement of special cortical mechanisms for perceiving bodies (Downing et al., 2001; see also Tsao et al., 2003). Finally, the current evidence on perceptual inversion effects for cars in car experts is ambiguous: although both Gauthier et al. (2000) and Xu et al. (2002) found slightly larger differences in d′ for sequential matching of upright versus inverted cars in car experts than control subjects (d′ = 2.4 upright versus 1.6 inverted for experts, 1.4 upright versus 0.8 inverted for nonexpert controls, interaction not significant for Gauthier et al.; 2.9 upright versus 2.0 inverted for experts, 1.2 upright versus 0.7 inverted for controls, interaction significant for Xu et al.), scaling differences make the interpretation of these data difficult, and indeed ratios of upright to inverted d′ go in the opposite direction in both studies (larger inversion effects for nonexperts than experts). It would clearly be useful to resolve these many ambiguities by measuring perceptual inversion effects in real-world experts when scaling effects are eliminated.

Finally, the one study that included a direct behavioral test of configural processing on real-world experts[3] found that objects-of-expertise are *not* processed configurally. Tanaka et al. (1996; cited in Tanaka & Gauthier, 1997) used the part-whole effect and found no influence of expertise (a minimum of 5 years, and mostly more than 10 years) for cars, biological cells, and Rottweiler dogs. Specifically, for each object class, a small whole-part advantage was of a similar size in experts (mean = 8.0 percentage points) and novices (8.3 percentage points) and both groups showed a far smaller effect for objects than for faces (23 percentage points).

Summary of Behavioral Studies

Currently, there are no behavioral data that unequivocally demonstrate "facelike" configural processing for objects-of-expertise. The failure to find such effects in laboratory-trained greeble experts does not seriously damage the expertise hypothesis, in that 10 hrs of training might not be expected to produce sufficient expertise. Yet, even in real-world experts, the one study that tested directly for configural processing failed to find any configural effect in the part-whole paradigm (Tanaka et al., 1996, cited in Tanaka & Gauthier, 1997). This leaves large inversion effects on memory tasks as the only result suggestive of an expertise effect (and, as discussed earlier, this could result from properties of long-term memory rather than properties of object processing). Most important, even clear behavioral evidence that objects of expertise are processed like faces would leave open the important question of whether this similarity reflects engagement of the very same mechanisms by faces and by objects of expertise, or the engagement of distinct mechanisms with similar functional properties. Neural measures are well suited to distinguish these alternatives.

Neural Evidence

How Faces Are Special in Neural Processing: The FFA, N170, and M170

Several face-selective neural responses have been reported based on research using fMRI, ERPs, and MEG in ordinary people—that is, those not expert in any particular object domain. Most relevant to the expertise debate are the fusiform face area (FFA) in fMRI studies, and the N170 and M170 responses measured with event-related potentials (ERPs) and magnetoencephalography (MEG), respectively.

The FFA is a cortical region in the fusiform gyrus that responds twice (or more) as strongly in fMRI when subjects view faces than when they view any other class of visual stimuli yet tested, even when the object task is more difficult than the face task, and even when all members of the object class (e.g., hands) share a basic configuration (Kanwisher, McDermott, & Chun, 1997; McCarthy, Puce, Gore, & Allison, 1997; Tong et al., 2000). Although the FFA shows only weak face inversion effects (Kanwisher, Tong, & Nakayama, 1998), the magnitude of the fMRI signal from this region is correlated on a trial-by-trial basis with successful identification of faces (Grill-Spector et al., 2004). Evidence that this region is not only activated during face recognition but is also critically involved in face recognition comes from the study of a patient with a very small lesion in this region who suffered profound prosopagnosia but no detectable deficit in object recognition (Wada & Yamamoto, 2001; see also Barton, Press, Keenan, & O'Connor, 2002; Puce, Allison, & McCarthy, 1999; Mundel et al., 2003)

The *N170* is an ERP component occurring about 170 ms after stimulus onset over occipitotemporal sensors (Bentin et al., 1996; Jeffreys, 1996). Its amplitude at many

scalp sensors is larger for faces than for a wide range of other objects (Carmel & Bentin, 2002). A face-specific inversion effect has been reported for the N170: for faces but not other objects, the N170 is delayed by 10 ms for the inverted orientation compared to upright (Rossion et al., 1999; Rossion et al., 2000). Similar properties have been reported for the face-selective "M170" response recorded with MEG (Liu, Higuchi, Marantz, & Kanwisher, 2000). It is not yet clear whether the N170 and M170 arise from the same cortical source, or whether the source for either is the FFA.

Neural Findings: Laboratory-Trained Experts

Gauthier et al. (1999) scanned subjects looking at faces and greebles. They found that activation for upright minus inverted greebles in the FFA region increased throughout greeble training. While Gauthier et al. interpreted their data as evidence for an expertise effect in the FFA, there are several problems with this conclusion. First, rather than measure the percent signal change from baseline for each stimulus type, they reported only the difference between upright and inverted orientations; this tells us nothing about the crucial question of the magnitude of response to upright greebles and upright faces after training. (This important distinction is easily lost; for example, Bentin and Carmel (2002) wrongly describe the Gauthier et al. result as "greebles recruited the FFA to nearly the same degree as faces do"). Second, the fMRI inversion effect is an odd choice as a marker of faceline processing because it is found only very weakly (Kanwisher, Tong, & Nakayama, 1998) or not at all (Haxby et al., 1999) for faces in the FFA. Third, the "FFA" was defined as a large square region of interest, over a cm on a side, a method that virtually guarantees the inclusion of voxels neighboring but not in the FFA. Finally, the "activation" was defined as the sum across the 64 voxels in this ROI of t values resulting from a comparison of upright to inverted responses within each voxel (after excluding all t values less than 0.1). This truncated "sum-of-ts" measure (see also Gauthier & Tarr, 2002) confounds an increase in signal change for upright stimuli after training with a reduction in variance. These problems leave the results difficult to interpret.

In an N170 study, Rossion and colleagues (2002) examined whether greeble training induced a facelike pattern of inversion effects. In terms of amplitude, greeble experts showed the predicted direction of effect in the left hemisphere (inverted minus upright greebles = 0.49 microvolts), but there was a similar-sized effect in the reverse-to-predicted direction in the right hemisphere (–0.50 microvolts). With respect to latency, there was a significant interaction of orientation (upright versus inverted) by session (pretraining versus posttraining) for greebles but not for faces. However, the effect for greebles (the latency delay for inverted versus upright stimuli) after training was significant only in the left hemisphere (5 msecs), not in the right (1.6 msecs; cf. 10 msecs for faces). Because no evidence was presented that the N170 recorded in this experiment was face selective, this study does not address the critical question of

whether the same neural mechanisms are engaged in processing faces and objects of expertise. Indeed, given that behavioral, fMRI, and ERP markers for face processing are all right lateralized, the left-lateralized expertise effects in this study would seem if anything to provide evidence for a dissociation, not an association, between expertise and face processing.

Neural Findings: Real-World Experts

Words are probably the only stimulus class for which we have perceptual expertise that approaches our expertise for faces. Words produce very weak FFA responses (Puce et al., 1996), demonstrating that perceptual expertise alone is insufficient to strongly engage the FFA. However, words and letters do not share a first order configuration, and it remains possible that expertise has an effect only where this is the case.

Gauthier et al. (2000) reported greater FFA activation for cars and birds than control objects in car and bird experts (19 years experience), respectively. This result has been confirmed in an event related design (Xu & Kanwisher, 2001). Note, however, that the expertise effect is small and percent signal increase from fixation remains twice as large for faces as for cars in car experts in both studies. Further, although Gauthier et al. emphasize as their strongest finding the correlation across subjects between behavioral expertise for cars/birds and the FFA response to cars/birds, this was found only for fMRI data collected during performance of a location discrimination task (on objects of expertise); it is not clear why they found no such correlation between behavioral expertise and fMRI responses to objects of expertise during a task requiring discrimination of objects of expertise. In a third study, Rhodes et al. (2004) scanned lepidoptera experts and found little overlap between the regions in the fusiform gyrus that were activated by lepidoptera (versus objects) and those activated by faces (versus objects); they concluded that distinct neural populations are tuned to the two object classes. Finally, Grill-Spector, Knouf, & Kanwisher (2004) found a clear trial-by-trial correlation between success at individual face recognition and the magnitude of the rFFA response to faces in car experts, but no trial-by-trial correlation between subordinate-level identification of cars in the same subjects. These data suggest that the weak engagement of the rFFA by objects of expertise is not causally related to expert object identification.

In terms of the N170, Tanaka and Curran (2001) reported a higher N170 response for bird than dog stimuli in bird experts, and vice versa for dog experts (all had at least 10 years of experience, most had 20 years). However, the scalp location where these expertise effects were found was different from the usual site of the face-selective N170, and no evidence was presented that these sites produce face-selective N170s. Thus, there is no reason to think that the expertise effects on the N170 reported in this study reflect the same neural source as the face-selective N170. Another general problem was that the dogs and birds used in the Tanaka and Curran (2001) study had

faces, and the N170 could reflect a response to these faces (enhanced by expertise) rather than generic expertise (Bentin & Carmel, 2002).

Another study of effects of expertise for cars on the N170 (Gauthier et al., 2003) reports a number of complex high-order interactions that are argued to be consistent with the expertise hypothesis, but fails report the results of the more straightforward prediction that the response to cars relative to control stimuli should be higher for experts than novices in face-selective sensors. Such an effect may be present in figure 3a of that paper, but this effect appears to be present only in the left hemisphere, not the right. Further, the paper mentions that different scalp distributions were found for the responses to cars and faces, suggesting a dissociation, not an association, between the processing of the two stimulus classes.

The only magnetoelectrophysiological study that actually tested the key prediction of the Expertise hypothesis—that the same neural response should show both face selectivity and enhanced responses for objects of expertise—decisively refuted the hypothesis. In particular, Xu, Liu, and Kanwisher (2004) found that the face-selective M170 was no greater for cars (relative to objects) in car experts than in control subjects; they also found that whereas the M170 amplitude was correlated trial-by-trial with successful face recognition, it was not correlated with successful car recognition.

Summary of Neural Processing Studies

Currently, the only real evidence of an expertise effect on either the FFA, N170, or M170 comes from the small but significant increase in the FFA response to objects-of-expertise in real world experts. According to the expertise hypothesis, the fact that car experts still show twice the FFA response to faces as to cars would be attributed to car experts being less experienced with cars than with faces. While this is one explanation, there are others. For example, because attention enhances fMRI activation (Wojciulik, Kanwisher, & Driver, 1998), the expertise effect could reflect greater attentional engagement by experts than novices. Even a task requirement to attend to identity (Gauthier, Skudlarski, Gore, & Anderson, 2000), or the use of an event related rather than blocked design (Xu & Kanwisher, 2001), may not completely override the greater interest in stimuli from the expert domain. The fact that ERP studies apparently show effects of expertise lasting hundreds of milliseconds after the N170 (e.g., figure 2 in Tanaka & Curran, 2001; figure 3a in Gauthier et al., 2003) is consistent with an attentional account.

Conclusions

We have examined in detail findings cited as support for the claim that facelike special processing emerges with experience for objects-of-expertise. We have shown that the empirical evidence *for* the expertise hypothesis is currently not strong, being essen-

tially limited to a large inversion effect on memory for dog experts (which could be attributed to memory rather than object recognition processes) and a larger-than-other-objects (but much smaller than faces) FFA activation effect in car and bird experts (which could be attributed to attentional differences). We have also reviewed evidence *against* the expertise hypothesis, namely (1) the lack of any configural processing in the behavioral greeble studies and, more strongly, (2) the lack of any effect of real-world expertise in the part-whole paradigm, as well as (3) the lack of any effect of expertise on the face-selective M170, and (4) the lack of a correlation across subjects between rFFA activation during identification of objects of expertise and degree of expertise (Gauthier et al., 2000) as well as the lack of a correlation across trials between rFFA activation and successful car identification in car experts (Grill-Spector et al., 2004). Although all of the evidence discussed so far tests the hypothesized identity between face processing and expertise by considering the case of nonface expertise, the syndrome of congenital prosopagnosia enables us to consider the converse case of nonexpert face processing. Although the existence of apparently normal FFAs in these subjects (Hasson et al., 2003; see also Rossion et al., 2003), who have apparently never been experts at face processing at any point in their lives, indicates that the FFA is not sufficient for face recognition, it also shows that expertise is not necessary for FFA activation.[4] (In contrast, patient evidence suggests that the FFA is necessary for face recognition; Wada & Yamamoto 2001 and Barton et al., 2002.) Finally, perhaps the strongest evidence against the expertise hypothesis comes from the double dissociation in the neuropsychology literature in which some patients are impaired at face recognition but not at nonface identification of objects of expertise (Sergent & Signoret, 1992), while others show the opposite pattern (Moscovitch, Winocur, & Behrmann, 1997).

While we suspect that the alternative face specificity hypothesis will turn out to be correct one, we note that this conclusion is at present somewhat open. Quite a few studies of expertise have been published, but many have used laboratory-trained greeble experts, and relatively few have tested real-world experts with many years of experience. Moreover, key behavioral markers for "face-specific" processing are yet to be been tested in real-world experts. This includes the composite effect, and relational versus local alteration paradigms (where relational alterations to faces are sometimes more strongly affected by inversion than local part alterations in many studies; e.g., Leder & Bruce, 1998). Should these paradigms turn out to show evidence of configural processing for objects-of-expertise, the face specificity hypothesis would need to be reconsidered. Similarly, any clear evidence that the FFA itself (as opposed to nearby cortical regions) is necessary for expert identification of nonfaces would seriously challenge the face specificity hypothesis.

The research reviewed above suffers from two important shortcomings. First, even real-world experts are never as expert in any other stimulus domain as they are with

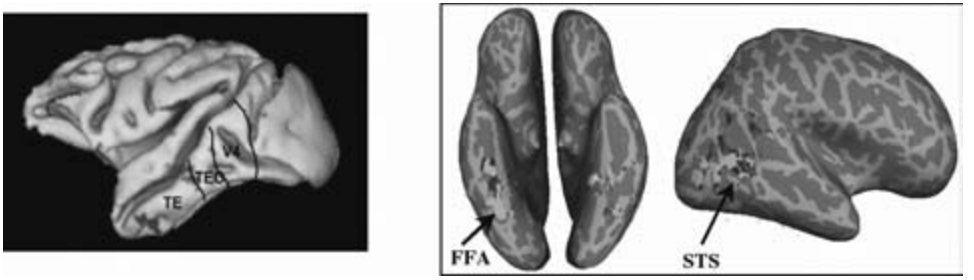


Figure 17.3
Regions responding more strongly in fMRI during viewing of faces than objects in a macaque (*left*, adapted from Tsao et al., 2003) and in an inflated human brain (*right*). It is unclear whether the macaque face area is homologous to the fusiform face area or the face-selective region in the STS region. See plate 20 for color version.

faces, so expertise effects that are much weaker than the corresponding effects for faces can generally be excused as reflecting insufficient expertise with the nonface domain. Second, because disruption methods are very limited in humans, it is difficult to test the *necessity* of face-selective cortical regions for processing objects of expertise. A possible new avenue for addressing both of these problems comes from the recent report (Tsao et al., 2003) of face areas in macaques (figure 17.3, plate 20), candidate homologs to the human FFA. To further investigate the homology across species, it will be important to test whether monkeys show the same behavioral signatures of configural processing of faces that have been so extensively demonstrated in humans. If it turns out that monkeys and humans process faces in similar ways, then further research in monkeys will make possible strong tests of the expertise hypothesis that avoid the shortcomings of the human literature. First, disruption methods not available in human research could be used to test the necessity of face-selective cortex for processing objects of expertise in monkeys. Second, by controlling perceptual experience from birth, it should be possible for the first time to conduct strong behavioral and neural tests of the expertise hypothesis in animals whose perceptual experience with faces and with other stimuli is matched.

Acknowledgments

Preparation of this article supported by Australian Research Council–Discovery grant DP0208630 to EM and NEI grant EY13455 to NK. We thank Marlene Behrmann, Brad Duchaine, Jia Liu, Morris Moscovitch, Rachel Robbins, and Galit Yovel for comments.

Notes

1. Note that, in some cases in this section, values have been estimated from figures.

2. The misorientation effect did disappear more slowly with practice in experts than in novices, but this is not a test for a "face-specific" effect: in normal nonexpert object recognition, misorientation effects for objects with a highly overlearned upright bias (letters, chairs, penguins, etc.) take 3–30 times through the stimulus set to disappear (Jolicoeur, 1985; McKone & Grenfell, 1999).

3. Gauthier, Curran, Curby, and Collins (2003) claimed that holistic processing was greater in car experts than in controls. They used a method bearing some similarity to the composite paradigm, but did not make the usual comparison between aligned and misaligned stimuli (which are matched for response competition from the other half; see figure 17.2). Instead, they defined "holistic processing" as the performance decrement when the to-be-ignored half suggested a response inconsistent, as opposed to consistent, with that required to the target half. This definition as merely the inability to ignore a notionally irrelevant component of the stimulus display is not what is usually meant by "holistic processing": indeed, under this definition, the Stroop effect (i.e., the difficulty of ignoring the word "red" when trying to name the color of the ink it is printed in) would incorrectly be interpreted as showing that color and word identity are processed together "holistically."

4. We thank Rebecca Saxe for pointing out the relevance of congenital prosopagnosia to the expertise hypothesis.

References

Bartlett, J. C., & Searcy, J. (1993). Inversion and configuration of faces. *Cognitive Psychology*, *25*, 281–316.

Barton, J. J., Press, D. Z., Keenan, J. P., & O'Connor, M. (2002). Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology*, *58*, 71–78.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of human face perceptions in humans. *Journal of Cognitive Neuroscience*, *8*, 551–565.

Bentin, S., & Carmel, D. (2002). Accounts for the N170 face-effect: A reply to Rossion, Curran and Gauthier. *Cognition*, *85*, 197–202.

Carmel, D., & Bentin, S. (2002). Domain specificity versus expertise: Factors influencing distinct processing of faces. *Cognition*, *83,* 1–29.

de Haan, M., Humphreys, K., & Johnson, M. H. (2002). Developing a brain specialized for face perception: A converging methods approach. *Developmental Psychobiology*, *40*, 200–212.

de Renzi, E. (1986). Current issues in prosopagnosia. In H. D. Ellis, M. A. Jeeves, F. Newcombe, & A. W. Young (Eds.), *Aspects of Face Processing* (pp. 243–252). Martinus Nijhiff: Dordrecht.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. J. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neurosciences*, *4*, 2051–2062.

Diamond, R., & Carey, S. (1986). Why faces are not special: An effect of expertise. *Journal of Experimental Psychology*: *General*, *115*, 107–117.

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, *293*(5539), 2470–2473.

Farah, M. J., Rabinowitz, C., Quinn, G. E., & Liu, G. T. (2000). Early commitment of neural substrates for face recognition. *Cognitive Neuropsychology*, *17*, 117–123.

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience*, *6*, 428–432.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, *3*, 191–197.

Gauthier, I., & Tarr, M. J. (1997). Becoming a "greeble" expert: Exploring mechanisms for face recognition. *Vision Research*, *37*, 1673–1682.

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior. *Journal of Experimental Psychology*: *Human Perception and Performance*, *28*, 431–446.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform "face area" increases with expertise in recognizing novel objects. *Nature Neuroscience*, *2*, 568–573.

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training "greeble" experts: A framework for studying expert object recognition processes. *Vision Research*, *38*, 2401–2428.

Grill-Spector, K. (2003). The functional organization of the ventral visual pathway and its relationship to object recognition. In N. Kanwisher & J. Duncan, (Eds.), *Attention & Performance XX*: *Functional Neuroimaging of Visual Cognition*. London: Oxford University Press.

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The Fusiform face area subserves face detection and face identification, not generic within-category identification. *Nature Neuroscience*, *7*, 555–562.

Hasson, U., Avidan, G., Deouell, L. Y., Bentin, S., & Malach, R. (2003). Face-selective activation in a congenital prosopagnosic subject. *Journal of Cognitive Neuroscience*, *15*, 419–431.

Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., & Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. *Neuron*, *22*, 189–99.

Jeffreys, D. A. (1996). Evoked potential studies of face and object processing. *Visual Cognition*, *3*, 1–38.

Jolicoeur, P. (1985). The time to name disoriented objects. *Memory and Cognition*, *13*, 289–303.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4302–4311.

Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition*, *68*, B1–B11.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Early visual experience and face processing. *Nature*, *412*, 786.

Leder, H., & Bruce, V. (1998). Local and relational aspects of face distinctiveness. *Quarterly Journal of Experimental Psychology*, *51A*, 449–473.

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: An MEG study. *Nature Neuroscience*, *5*, 910–916.

Liu, J., Higuchi, M., Marantz, A., & Kanwisher, N. (2000). Magnetophysiological studies of face and non-face processing. *NeuroReport*, 11, 337–341.

Maurer, D., Le Grand, R., & Mondloch, J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6,* 255–260.

McCarthy, G., Puce, A., Gore, J., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, *9*, 605–610.

McKone, E., & Grenfell, T. (1999). Orientation invariance in naming rotated objects: Individual differences and repetition priming. *Perception and Psychophysics*, *61*, 1590–1603.

McNeil, J. E., & Warrington, E. K. (1993). Prosopagnosia: A face-specific disorder. *Quarterly Journal of Experimental Psychology*, *46A*, 1–10.

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLEARN: A two process theory of infant face recognition. *Psychological Review*, *98*, 164–181.

Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience*, *9*, 555–604.

Mundel, T., Milton, J. G., Dimitrov, A., Wilson, H. W., Pelizzari, C., Uftring, S., et al. (2003). Transient inability to distinguish between faces: electrophysiologic studies. *Journal of Clinical Neurophysiology*, *20*, 102–110.

Nederhouser, M., Mangini, M., Biederman, I., & Okada, K. (2002). Matching face-like objects is invariant to differences in direction of contrast. Paper presented at the 10th Annual Workshop on Object Perception and Memory, Kansas City, MO.

Puce, A., Allison, T., Asgari, M., Gore, J., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letter strings, and textures: A functional magnetic resonance imaging study. *Journal of Neuroscience*, *16*, 5205–5215.

Puce, A., Allison, T., & McCarthy, G. (1999). Electrophysiological studies of human face perception. III: Effects of top-down processing on face-specific potentials. *Cerebral Cortex*, *9*, 445–458.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, *81*, 275–280.

Rhodes, G. (1988). Looking at faces: First-order and second-order features as determinants of facial appearance. *Perception*, *17*, 43–63.

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition*, *47*, 25–57.

Rhodes, G., Byatt, G., Michie, P. T., & Puce, A. (2004). Is the fusiform face area specialized for faces, individuation, or expert individuation? *Journal of Cognitive Neuroscience*, *16*, 189–203.

Rossion, B., Delvenne, J. F., Debatisse, D., Goffaux, V., Bruyer, R., Grommelinck, M., et al. (1999). Spatio-temporal localization of the face inversion effect: an event related potentials study. *Biological Psychology*, *50*, 173–189.

Rossion, B., Gauthier, I., Goffaux, V., Tarr, M. J., & Crommelinck, M. (2002). Expertise training with novel objects leads to left-lateralized facelike electrophysiological responses. *Psychological Science*, *13*, 250–257.

Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., et al. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: An electrophysiological account of face-specific processes in the human brain. *Neuro-Report*, *11*, 69–74.

Scapinello, K. F., & Yarmey, A. D. (1970). The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. *Psychonomic Science*, *21*, 329–330.

Schneider, W. X., Korkel, J., & Weinert, F. (1989). Domain specific knowledge and memory performance: A comparison of high and low aptitude children. *Journal of Educational Psychology*, *81*, 301–312.

Sergent, J., & Signoret, J. L. (1992). Varieties of functional deficits in prosopagnosia. *Cerebral Cortex*, *2,* 375–388.

Subramanian, S., & Biederman, I. (1997). Does contrast reversal affect object identification? *Investigative Ophthalmological and Visual Sciences*, *38*, 998.

Tanaka, J., & Gauthier, I. (1997). Expertise in object and face recognition. *The Psychology of Learning and Motivation*, *36*, 83–125.

Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological Science*, *12*, 43–47.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, *46A*, 225–245.

Tanaka, J. W., Kay, J. B., Grinnell, E., Stansfield, B., & Szechter, L. (1996). Face recognition in young children: When the whole is greater than the sum of its parts. *Visual Cognition*, *5*, 479–496.

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory and Cognition*, *25*, 583–592.

Tarr, M. J. (2003). Visual object recognition: Can a single mechanism suffice? In M. Peterson & G. Rhodes (Eds.), *Perception of Faces, Objects and Scenes: Analytic and Holistic Processes* (pp. 172–211). Oxford University Press: London.

Tarr, M. J., & Cheng, Y. D. (2003). Learning to see faces and objects. *Trends in Cognitive Sciences, 7*, 23–30.

Tong, F., Nakayama, K., Moscovitch, M., Weinrib, O., & Kanwisher, N. (2000). Response properties of the human fusiform face area. *Cognitive Neuropsychology, 17*, 257–279.

Tsao, D. Y., Freiwald, W. A., Knutsen, T. A., Mandeville, J. B., & Tootell, R. B. (2003). Faces and objects in macaque cerebral cortex. *Nature Neurosci*ence, *6*, 989–995.

Wada, Y., & Yamamoto, T. (2001). Selective impairment of facial recognition due to a haematoma restricted to the right fusiform and lateral occipital region. *Journal of Neurological and Neurosurgical Psychiatry, 71*, 254–257.

Wheeler, D. D. (1970). Processes of word recognition. *Cognitive Psychology, 1*, 59–85.

Wojciulik, E., Kanwisher, N., & Driver, J. (1998). Covert visual attention modulates face-specific activity in the human fusiform gyrus: A fMRI study. *Journal of Neurophysiology, 79*, 1574–1478.

Xu, Y., & Kanwisher, N. (2001). What is the magnitude and theoretical significance of the FFA response to expert stimuli? Paper presented at the Society for Neuroscience, San Diego, CA.

Xu, Y., Liu, J., & Kanwisher, N. (2004). The M170 is selective for faces, not for expertise. *Neuropsychologia*, in press.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*, 141–145.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*, 747–759.

18 Representation of Object Images by Combinations of Visual Features in the Macaque Inferotemporal Cortex

Manabu Tanifuji, Kazushige Tsunoda, and Yukako Yamane

The primate ventral visual pathway plays a role in extracting features from object images that are necessary for object recognition. Inferior temporal (IT) cortex in monkeys is the final stage of the ventral visual pathway, and visual responses of neurons in this area have been extensively investigated to find these features extracted from object images. Area TE in IT cortex has been of particular interest, since this area is the final purely visual area in the ventral visual pathway (Desimone et al., 1984; Tanaka, Saito, Fukado, & Moriya, 1991; Kobatake & Tanaka, 1994). As described in detail below, these investigations suggest that an object image is not represented by a single feature but by a set of features. Thus, it is difficult to explore the sets of features necessary for object recognition by conventional single cellular recordings because these techniques only provide responses of a small number of neurons at the same time. The combination study of a optical imaging technique and single cellular recordings overcomes this difficulty and reveals how combination of features could specifically represent object images in area TE.

Structural Description and View-Based Object Representation

Object representation in our brain is also an issue in computational and psychological studies of object vision. One proposal in these fields is that objects are represented by parts and spatial relationships among parts. Multiple groups have investigated possible models based on this proposal, referred to as "structural description" (Biederman, 1987; Marr & Nishihara, 1978). According to this proposal, an object image is first decomposed into parts and then the combination of these parts and their spatial relationships is used for recognition. Thus, object recognition could be view independent unless object parts are occluded. It should be emphasized that representation of the spatial relationships as well as that of parts themselves is essential for recognition in these studies. On the other hand, another group of studies has proposed that recognition is based on object views (Poggio & Edelman, 1990; Tarr & Bulthoff, 1998; Ullman, 1998; Riesenhuber & Poggio, 1999). In latter models, a view of an object could

be directly compared with an exemplar view in memory without decomposing the object into parts. Explicit representation of object parts is not necessarily required in these models (Poggio & Edelman, 1990). However, appropriate codes that represent characteristic features of object images under particular viewing conditions could generate representation of object views with smaller cost and strengthen recognition capability (Tarr & Bulthoff, 1998; Riesenhuber & Poggio, 1999). These two dimensional features could be visual features that are less complex than the object images or parts of object images. Altogether, in both proposals, namely object representation based on structural description and view based representation, it is necessary to explore whether the brain extracts certain features or not, and if so, we have to consider what they are. Physiological investigations with the monkey inferior temporal cortex could provide concrete clues to consider these issues.

Evidence for the Representation of Visual Features by IT Neurons

There are a number of studies to search for features essential for activating these neurons using a broad variety of visual stimuli in area TE. These studies have shown that there are neurons responding equally well to object images and to visual features that are geometrically less complex than the object images (Desimone et al., 1984; Tanaka et al., 1991; Kobatake & Tanaka, 1994). In particular, Tanaka and colleagues examined responses of neurons in area TE with systematically simplified visual stimuli and found that visual stimuli sufficient to activate many neurons were visual features that are less complex than object images ("critical feature") (Kobatake & Tanaka, 1994) (figure 18.1). First, for each cell, they searched for the most effective stimulus among more than 100 three-dimensional object stimuli, including stuffed animals, plastic fruits and vegetables, and experimenter's hand and body. Then, they generated modifications of the most effective stimulus and examined responses evoked by the simplified stimuli. If the cell responded to one of the simplified stimuli equally well as compared to that to the original object stimulus, this new stimulus was further simplified. This procedure was repeated until the experimenter encountered a drastic decrease in responses due to simplification. Many of these features are combinations of simple shapes, colors, luminance gradient/contrast, and textures. These features are more complex than the optimal visual stimuli for cells in areas V1, V2, and V4, but still less complex than natural objects (figure 18.1). Figure 18.2 shows an example of the stimulus simplification procedure, where we found a combination of the circle and rectangle was essential for the maximal activation of the cell. Since TE neurons respond to visual features less complex than objects, representation specific to particular object images requires activities of multiple neurons in area TE. The functional imaging technologies such as intrinsic signal imaging provide an opportunity to investigate activation of multiple cells by object images.

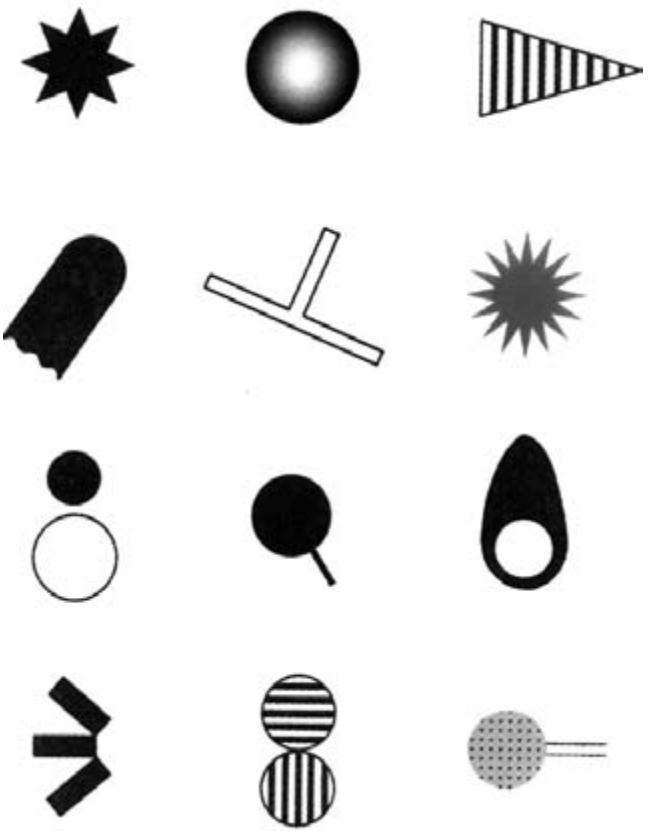


Figure 18.1
Representative critical features determined for IT cells. Twelve visual features are determined for different TE cells by stimulus simplification. (Adapted from Tanaka, 1993.)

Intrinsic Signal Imaging

Neurons with similar response properties are clustered into a column in area TE (Fujita, Tanaka, Ito, & Cheng, 1992). Thus, intrinsic signal imaging of columnar activation can be used to investigate spatial patterns of activation (Tsunoda, Yamane, Nishikazi, & Tanifuji, 2001; Wang, Tanaka, & Tanifuji, 1996; Wang, Tanifuji, & Tanaka, 1998). This technique measures the decrease in the degree of light reflection elicited by neural activation from the exposed cortical surface using a CCD camera (Grinvald et al., 1999). The reflection changes are due to metabolic changes elicited by neural activation including changes in deoxygenation of hemoglobin in capillaries (Grinvald et al., 1999). Although these reflection changes are not a direct measure of neural activation,

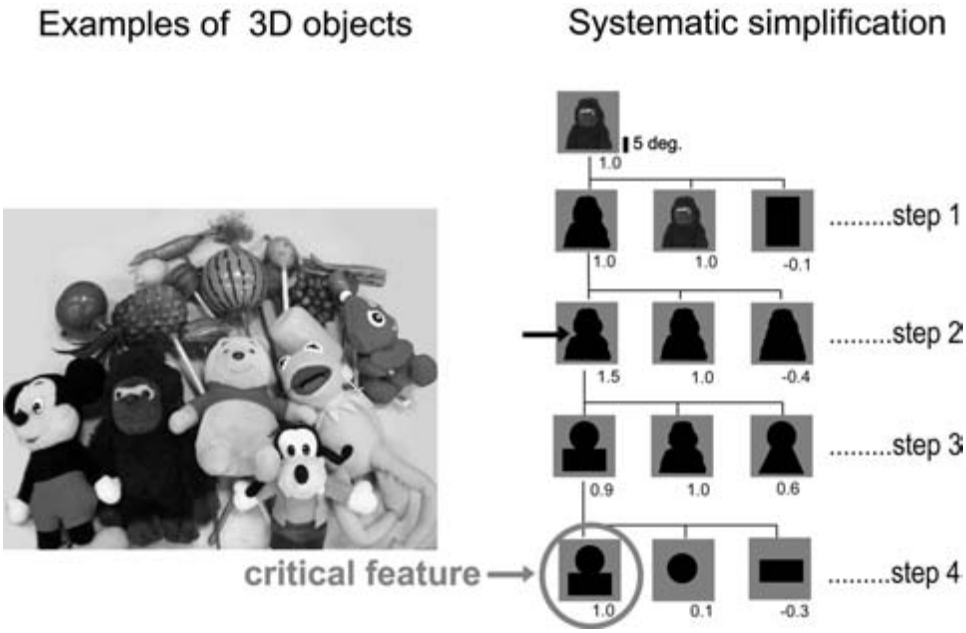


Figure 18.2

The "critical feature," the visual feature that maximally activates each cell, is determined by systematic stimulus simplification of the best object stimulus. First, we tested the cell with various 3D objects, including faces, hands, stuffed animals, plastic fruits and vegetables, and paper mounts (see the *left panel* for some examples). After determining the best stimulus, we simplified it step by step to find the simplest stimulus that maximally activates the cell (*right panel*). For example, at step 1 we compared the best colored object with its silhouette and found that the silhouette activated the cell equally well. The rightmost rectangle was taken as a control stimulus. The numbers below each picture indicate the response amplitudes normalized to the response to the reference stimulus, the best object. The stimulus that evoked at least more than 70 percent of the response elicited by the best stimulus in the previous step, was again examined in the next step as the reference stimulus. At step 2, we examined the effect of the "sharpness" of the corner at the junction of upper and lower parts (arrow) and found that the silhouette with the sharp corners was the most effective stimulus. From left to right, the stimuli were the silhouette with sharp corners, the silhouette that evoked the best response at the previous step, the silhouette without corners. Further simplification was carried out at step 3. Finally, we determined the critical feature as a combination of a circle and a rectangle because neither the upper nor lower part alone activated the cell. This particular cell was found in a spot with stimulus selectivity given in figure 18.6 (From Yamane and Tanifuji, unpublished observations.)

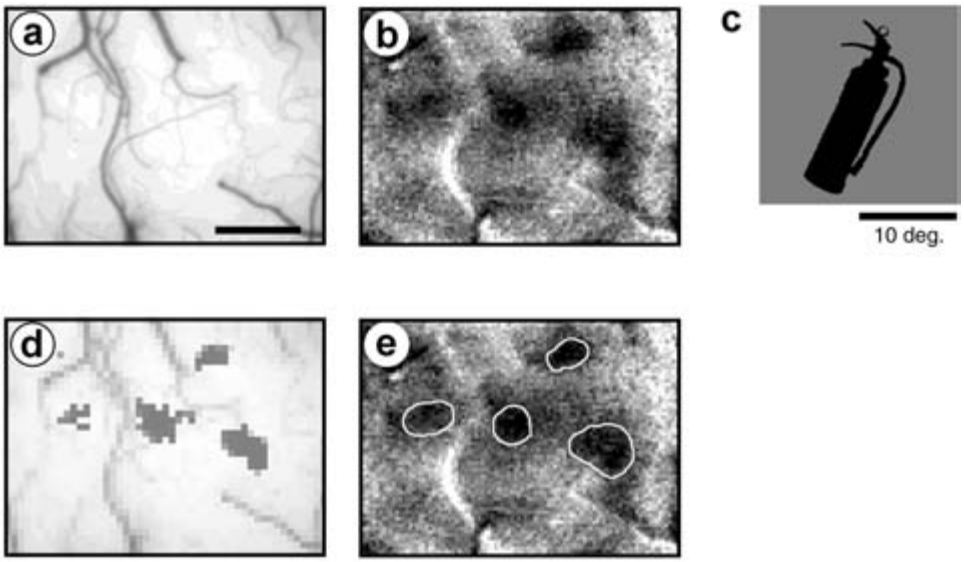


Figure 18.3
Intrinsic signal imaging detects local modulation of light absorption changes in area TE. (*a*) Portion of area TE where intrinsic signals were recorded. (*b*) A differential image showing a local increase in absorption by the visual stimulus shown in (*c*). (*d*) Active spots, where the degree of reflection change evoked by the stimulus was significantly greater than that without the stimulus presentation. The region with the highest significance level is in red, that with the lowest significant level in yellow ($p < 0.05$). (*e*) Active spots outlined by connecting pixels with 1/2 of the peak absorption value. (Modified from Tsunoda et al., 2001.) See plate 21 for color version.

intrinsic signals coincide well with the activity of neurons examined by conventional extracellular recordings (Wang et al., 1996, 1998; Tsunoda et al., 2001). Thus, if nearby neurons are simultaneously activated and elicit detectable changes in reflection, the spatial organization of the activated neurons can be investigated by this technique. Intrinsic signal imaging in area TE revealed multiple spots elicited by visual stimulation (figure 18.3, plate 21). These "active spots" could correspond to a column of cells with similar responsiveness in this area (Fujita et al., 1992).

Object Representations by Combined Activation of Neurons in Area TE

Intrinsic signal imaging revealed that complex objects activate multiple spots (figure 18.4a, plate 22). Each of these spots could represent a particular visual feature as proposed previously. To examine this idea, we compared distribution patterns of spots activated by a complex object with those activated by systematically simplified stimuli

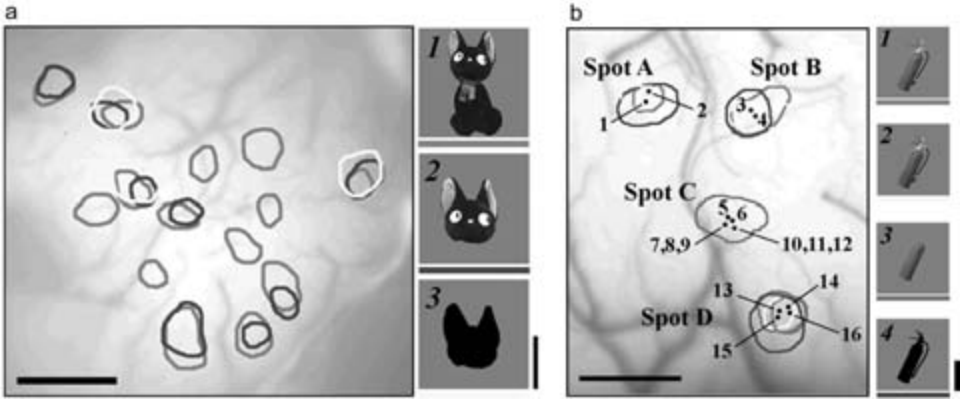


Figure 18.4
Representation of complex object images and their simplification in area TE (*a*) A case where simplified stimuli elicited only a subset of spots evoked by more complex stimuli. (*b*) A case in which new spots appeared when the original stimulus was simplified. The numbers (1 to 16) indicate electrode penetration sites. Horizontal scale bars: 1 mm; vertical scale bar: 10 deg. (Modified from Tsunoda et al., 2001.) See plate 22 for color version.

(Tsunoda et al., 2001). For example, we used a "black cat" (a-*1*) as the complex object image and then simplified it to its "head" (a-*2*), and to the "silhouette of its head" (a-*3*). The original image (a-*1*) elicited fourteen spots, but presenting the "head" (a-*2*) elicited only eight of the original fourteen spots. The silhouette (a-*3*) only activated three (yellow) of the eight spots elicited by the head (a-*2*). Thus, the simplified stimulus lacking part of features in the original image activates only a subset of the spots elicited by the stimuli before simplification. This result is consistent with the idea that individual columns represent visual features rather than object images in area TE.

Similarly, figure 18.4b shows that spots A and B disappeared but spot D remained when the "handle" and "hose" were removed from the original stimulus, "fire extinguisher." In addition to the disappearance of spots, this particular case shows that new spots can emerge by apparent simplification of an object: spot C appeared when the handle and hose were removed from the fire extinguisher. The emergence of spots by stimulus simplification suggests that part of the spots representing a visual feature were not activated if the feature was embedded in a complex object in some cases (see below). We have examined 12 pairs of activation patterns obtained before and after the simplification of the objects, and we observed changes in the distribution patterns consistent with either figure 18.4a or 18.4b for all of the pairs.

Visual Features Represented by Individual Spots

To determine critical features of these active spots, we recorded extracellular responses from 25 cells in the four spots shown in figure 18.4b, and analyzed the response properties of the cells in each spot (figure 18.5). The difference in optical response patterns to stimuli *1* and *3* in figure 18.4b suggests that spots A and B represented visual features related to the handle and hose of the fire extinguisher. Consistent results are obtained at the single cellular level: the handle and hose in isolation (a-*2* and b-*2*) as well as the silhouette of the original fire extinguisher (a-*1* and b-*1*) activates cells in spots A and B (figure 18.5). The cells in spot A were activated by the handle (figure 18.5, a-*3*) having protrusions, but not by the hose (figure 18.5, a-*4*). Furthermore, other stimuli with sharp protrusions, such as a "hand" (figure 18.5, a-*5*) and cat's head (figure 18.5, a-*6*), also activated the cells. Thus, the critical feature for the cells in spot A was "sharp protrusions." In contrast, cells in spot B were activated by the hose (figure 18.5, b-*4*), but neither by the handle (figure 18.5, b-*3*) nor a "line segment" (figure 18.5, b-*5*). Thus, the critical feature for the cells in spot B was an "asymmetric arc" (figure 18.5, b-4). The neural responses of cells in spots C and D were consistent with the imaging results in figure 18.4b: cells in spot C were activated by the "cylinder" but not by the original fire extinguisher (figure 18.5, c-*1* and *2*), and cells in spot D were significantly activated by both stimuli (figure 18.5, d-*1* and *2*). The critical feature for cells in spot D was a "rectangular shape" (figure 18.5, d-*3*), but cells also responded significantly to an "ellipse" (figure 18.5, d-*4*). Since there was no response to a "circle" (figure 18.5, d-*5*), we determined the critical feature of the spots as an "elongated structure".

Similarly, the simplest visual feature that could activate the cells in spot C was a rectangular shape (figure 18.5, c-*3*). In contrast to the cells in spot D, however, there was no activation by an ellipse (figure 18.5, c-*4*). In addition, the cells were inhibited by a circle (figure 18.5, c-*5*). Thus, these results suggest that the response properties of the cells in spot C (figure 18.4b) are determined by the balance between excitatory and inhibitory inputs: the excitatory inputs were given by a feature related to a rectangular shape and the inhibitory inputs are given by a feature related to a circle. This explanation would account for the lack of activation by the fire extinguisher, where the hose (circular shape) attached to the rectangular cylinder makes entire shape elliptical. These results suggest that some of the spots representing a particular feature are inactive when other features are presented together with that feature. This could explain the optical imaging results in which active spots appeared following simplification of the stimulus.

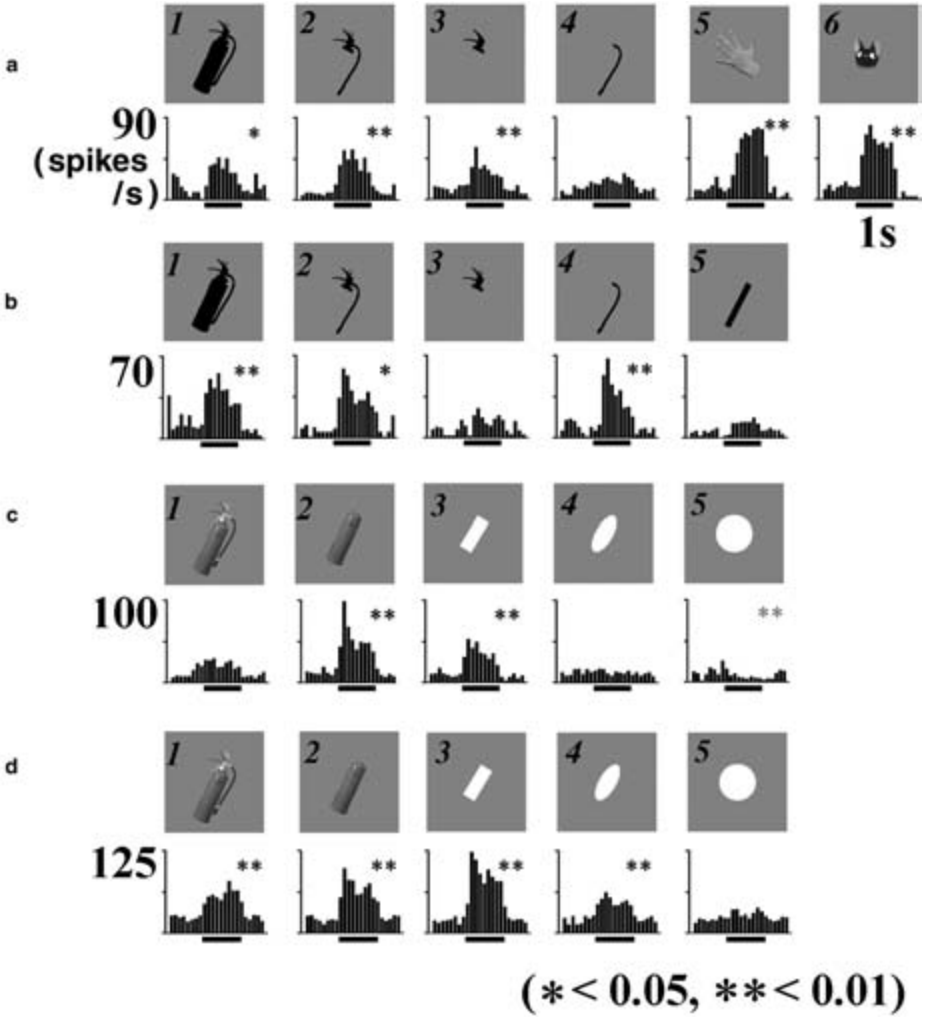


Figure 18.5

Visual responsiveness of representative cells in spots A–D in figure 18.4b. (*a*, *b*, *c*, and *d*) Responses in spots A (track 2, depth 620 μm), B (track 3, depth 540 μm), C (track 8, depth 280 μm) and D (track 16, depth 280 μm), respectively. Red asterisks indicate significant inhibition ($p < 0.01$). (Adapted from Tsunoda et al., 2001.)

Representation of Spatial Arrangement of Parts in Object Images

Examination of visual features represented by neurons in area TE suggested that at least some of the neurons in this area represent "local features" in object images, as neurons in spots A and B (figures 18.4, 18.5a, and 18.5b) represent "protrusions" and "asymmetric curvature," respectively. Since information about the spatial arrangement of "local features" is necessary for the specific representation of object images, some of the other spots may represent visual features related to the spatial arrangement of local features ("configurational information"). Here, we refer to "local features" as visual features that occupy part of an object image and are distinguishable from other parts of an object image by their particular shapes, colors, or textures. "Configurational information" is the information about the spatial relationship of "local features" themselves or about spatial relationship of parts including local features. To examine the representation of "configurational information," we investigated spots activated by an object (original, figure 18.6-1) and the same object with a gap introduced between parts of the object (figure 18.6-4), but not by a part alone (figures 18.6-2 and 18.6-3) (Yamane et al., 2001). These spots, if there were, do not simply represent local features in objects because either part is not essential for activation. Moreover, activation by the stimulus with an introduced gap indicates that local features appearing at the junction of two parts, such as sharp connecting corners in 18.6-1, are also not essential. In three monkeys, we indeed found some of active spots had stimulus selectivity described above. Extracellular recordings from cells in these spots showed that their critical features were combinations of vertically aligned two parts (figure 18.7a). In particular, the stimulus simplification procedure for these cells in this spot revealed that there was no activation by either part (for a representative case, see figure 18.2). These cells were less sensitive to color, texture, and local shapes of either part: (1) there were no changes in the responses after removing color and texture during the stimulus simplification procedure (figures 18.2 and 18.7a), (2) changes in the shapes of the parts did not significantly alter responses of these neurons, and (3) these cells responded equally well to object images even having different color, texture, and local shapes, as long as they had the global shape similar to the critical features (figure 18.7b). Thus, these neurons were not sensitive to the identity of parts (figure 18.8). In contrast, we found that these cells were highly selective to a particular spatial arrangement of the upper and lower parts (figure 18.9).

Summary and Conclusions

The combination of intrinsic signal imaging and extracellular recordings suggests that object images are represented as combinations of spots, and that each spot represents visual features less complex than the original object images (figures 18.4 and 18.5).

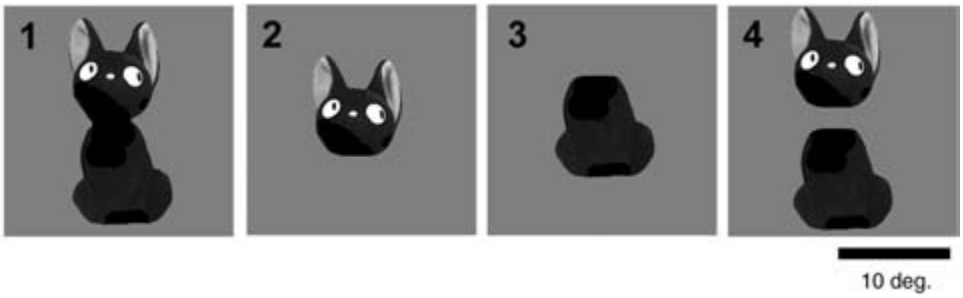


Figure 18.6
A representative set of visual stimuli used in intrinsic signal imaging for examination of the representation of the spatial arrangement of parts. The response properties shown in figures 18.1 and 18.7–18.9 were obtained from a spot activated by stimuli 1 and 4, but not by stimuli 2 and 3.

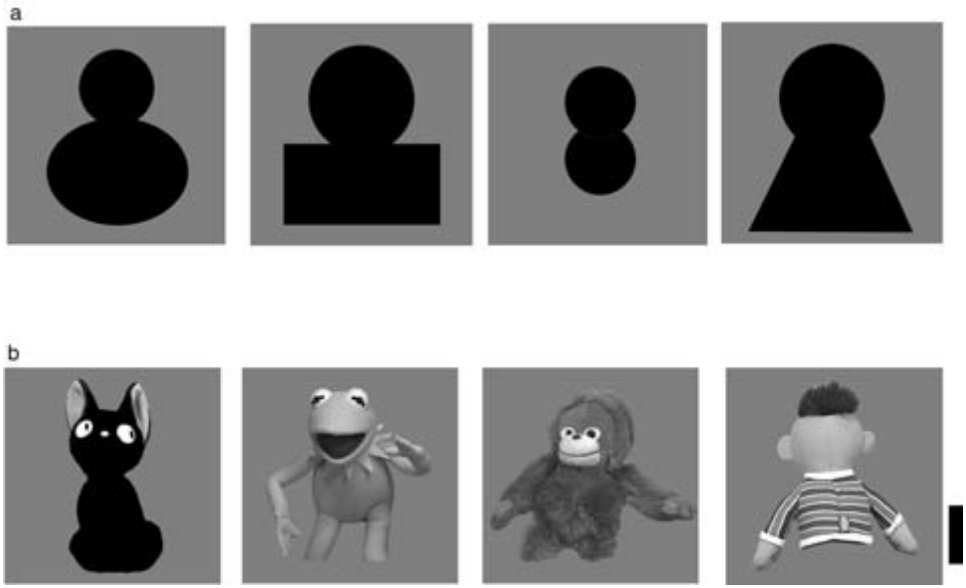


Figure 18.7
Effective stimuli for neurons in a spot identified by the stimuli in figure 18.6. (*a*) Representative critical features determined by stimulus simplification. Please note that, when color and texture are not essential, the stimulus was filled black (see figure 18.1) (*b*) The best object stimuli for these neurons, among 100 object stimuli examined before stimulus simplification. Scale bar: 5 deg.

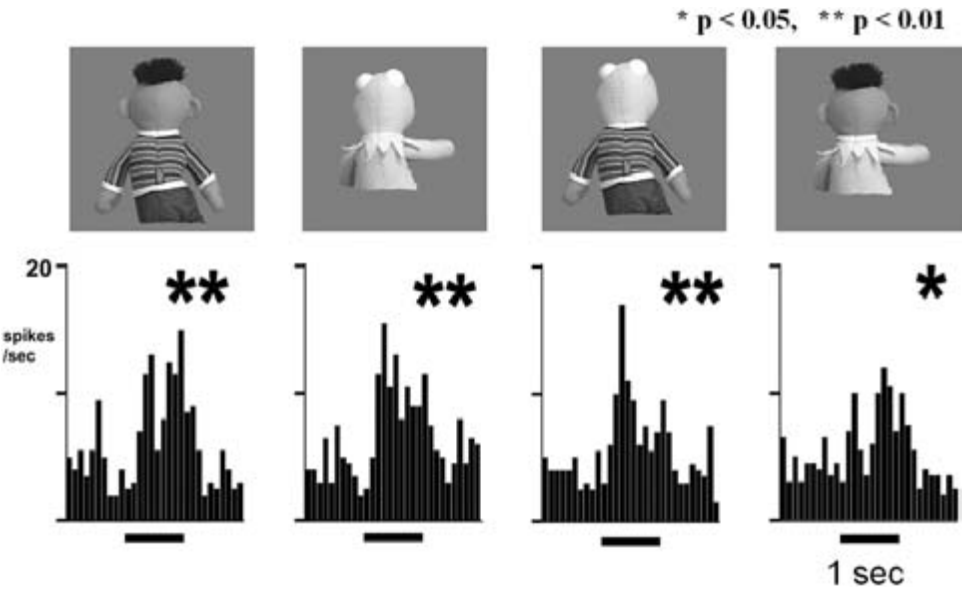


Figure 18.8
Responses of a neuron specific for particular spatial relationship among parts. The upper panel shows visual stimuli, and the lower panel indicates PSTHs showing responses of the neuron to the stimuli given above. The stimulus was presented for a 1-sec period, indicated by the horizontal line segment in each PSTH. These stimuli activated the cell equally well. *p < 0.05, **p < 0.01.

These visual features are common among various object images. For example, spot A in figure 18.4b represents "sharp protrusions," a common feature among fire extinguisher, hand, and cat (figure 18.5a). Thus, representation specific to object images requires combination of spots.

Spots do not necessarily represent "local features," but some of them represent visual features related to object configurations. We found that neurons in these spots responded to visual stimuli consisting of vertically aligned upper and lower part (figure 18.9), but were less selective to local features embedded in either part (figures 18.7 and 18.8). We consider that such neurons specify configuration of object images. Face neurons in area TE could play the same role. They respond specifically to a configuration specific for faces, but are less selective to individual faces (Desimone et al., 1984; Perrett et al., 1984; Baylis, Rolls, & Leonard, 1985; Yamane et al., 1988; Young & Yamane, 1992). A combination of neurons specific for local features and those specific for configurational information about object images such as spatial relationships among local features generates specific representation of object images in area TE.

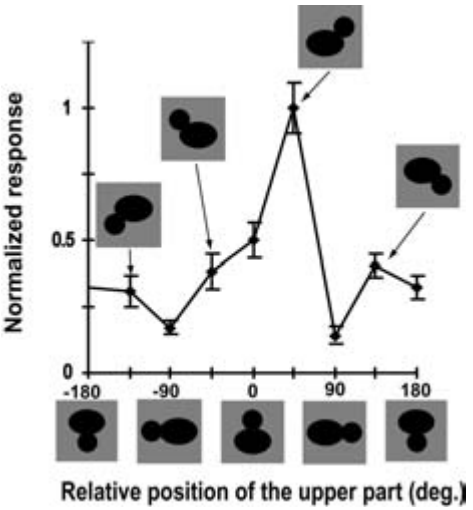


Figure 18.9
Response specificity of a representative cell to the spatial arrangement of parts. The upper part of the critical feature of the cell was rotated relative to the lower part. The horizontal axis indicates the angle between the line connecting the center of upper and lower parts of each stimulus and that of the critical feature. The vertical axis indicates normalized value of stimulus evoked responses. In this particular case, the best response was elicited by the stimulus with 45 deg, but many others respond maximally at 0 deg.

Finally, object-specific combination consists of active and inactive spots. For example, specific representation of the original fire extinguisher (figure 18.4b, stimulus b-1) requires spot C not to be activated; otherwise stimuli b-1 and b-3 could not be distinguished. Combinations of inactive as well as active spots increase the number of available activation patterns, and thus in general, could increase the number of objects to be specifically represented.

Earlier we described two frameworks for object representation. Based on our results that object images are represented with local visual features and spatial relationship among arbitrary local visual features, one may consider that our results support the models based on structural description. At present, however, we have to make a clear distinction between our results and those theoretical frameworks. First, parts are conventionally defined as the ones naturally distinguishable by discontinuities at the minima of negative curvature of the object shape. The visual features described here do not necessarily correspond to such parts. Second, one related argument among the computational models is whether object recognition is view-dependent or not. At present, we do not know whether the responses of neurons to the critical feature are invariant for different views or the features. Finally, the models based on view-based

recognition do not necessarily reject the intermediate representation of features (Riesenhuber & Poggio, 1999). Thus, the question whether the object representation is view-based or not remains to be clarified through future experiments. Particularly, it is necessary to clarify the tuning property of neurons to different views of visual features. The important conclusion obtained from our investigation is that, regardless to the theoretical frameworks, object images are represented in a distributed manner through local features and spatial relationships among them.

Acknowledgments

This work was partly supported by Research Fellowships of the Japan Society for the Promotion of Young Scientists to Y. Y.

References

Baylis, G. C., Rolls, E. T., & Leonard, C. M. (1985). Selectivity between faces in the responses of a population of neurons in the cortex in the superior temporal sulcus of the monkey. *Brain Research*, *342*, 91–102.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–147.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience*, *4*, 2051–2062.

Fujita, I., Tanaka, K., Ito, M., & Cheng, K. (1992). Columns for visual features of objects in monkey inferotemporal cortex. *Nature*, *360*, 343–346.

Grinvald, A., Shoham, D., Shmuel, A., Glaser, D., Vanzetta, I., Shtoyerman, E., Slovin, H., Wijnbergen, C., Hildesheim, R., & Arieli, A. (1999). In-vivo optical imaging of cortical architecrure and dynamics. In U. Windhorst & H. Johansson (Eds.), *Modern Techniques in Neuroscience Research* (pp. 893–970). Berlin: Springer.

Kobatake, E., & Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology*, *71*, 856–867.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London B., Biological Sciences*, *200*, 269–294.

Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1984). Neurons responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity and relation to perception. *Human Neurobiology*, *3*, 197–208.

Poggio, T., & Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature*, *343*(6255), 263–266.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*, 1019–1025.

Tanaka, K., Saito, H., Fukada, Y., & Moriya, M. (1991). Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *Journal of Neurophysiology*, *66*, 170–189.

Tarr, M. J., & Bulthoff, H. H. (1998). Image-based object recognition in man, monkey and machine. *Cognition*, *67*, 1–20.

Tsunoda, K., Yamane, Y., Nishizaki, M., & Tanifuji, M. (2001). Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. *Nature Neuroscience*, *4*, 832–838.

Ullman, S. (1998). Three-dimensional object recognition based on the combination of views. *Cognition*, *67*(1–2), 21–44.

Wang, G., Tanaka, K., & Tanifuji, M. (1996). Optical imaging of functional organization in the monkey inferotemporal cortex. *Science*, *272*, 1665–1668.

Wang, G., Tanifuji, M., & Tanaka, K. (1998). Functional architecture in monkey inferotemporal cortex revealed by in vivo optical imaging. *Neuroscience Research*, *32*, 33–46.

Yamane, S., Kaji, S., & Kawano, K. (1988). What facial features activate face neurons in the inferotemporal cortex of the monkey? *Experimental Brain Research*, *73*, 209–214.

Yamane, Y., Tsunoda, K., Matsumoto, M., Phillips, A., & Tanifuji, M. (2001). Decomposition of object images by feature columns in macaque inferotemporal cortex. *Society for Neuroscience Abstr 399.6*.

Young, M. P., & Yamane, S. (1992). Sparse population coding of faces in the inferotemporal cortex. *Science*, *256*, 1327–1331.

Contributors

Suliann Ben Ahmed Institut des Sciences Cognitives, Bron, France

Céline Amiez U371 Inserm, Bron, France

Nick Barraclough School of Psychology, University of St. Andrews, Fife, Scotland, United Kingdom

Elizabeth M. Brannon Center for Cognitive Neuroscience, Duke University, Durham, North Carolina

Giovanni Buccino Department of Neuroscience, School of Medicine, University of Parma, Italy

Jean-Pierre Changeux Collège de France and Pasteur Institute, Paris, France

Stanislas Dehaene Service Hospitalier Frédéric Joliot, Orsay, France

Jean-René Duhamel Institut des Sciences Cognitives, Bron, France

Juan-Carlos Gómez School of Psychology, University of St. Andrews, Fife, Scotland, United Kingdom

Marc D. Hauser Department of Psychology, Harvard University, Cambridge, Massachusetts

Jean-Jacques Hublin Department of Anthropology, University of Bordeaux, Talence, France

Atsushi Iriki Department of Maxillofacial Biology, Tokyo Medical and Dental University, Tokyo, Japan

Tjeerd Jellema Department of Experimental Psychology, Helmholtz Research Institute, Utrecht University, The Netherlands

Jean-Paul Joseph U371 Inserm, Bron, France

Nancy Kanwisher Department of Brain and Cognitive Sciences, Massachussets Institute of Technology, Cambridge, Massachusetts

Zoe Kourtzi Max Planck Institute for Biological Cybernetics, Tübingen, Germany

Erica N. Lorincz School of Psychology, University of St. Andrews, Fife, Scotland, United Kingdom

Nikos K. Logothetis Max Planck Institute for Biological Cybernetics, Tübingen, Germany

Giuseppe Luppino School of Medicine, University of Parma, Italy

Elinor McKone School of Psychology, Australian National University, Canberra, Australia

Earl K. Miller The Picower Center for Learning and Memory and Department of Brain and Cognitive Sciences, Massachussets Institute of Technology, Cambridge, Massachusetts

Andreas Nieder Primate NeuroCognition Laboratory, University of Tübingen, Germany

David I. Perret School of Psychology, University of St Andrews, Fife, Scotland, United Kingdom

Michael Petrides Department of Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, Montréal, Canada

Emmanuel Procyk U371 Inserm, Bron, France

Giacomo Rizzolatti Department of Neuroscience, School of Medicine, University of Parma, Italy

Jeffrey R. Stevens Department of Psychology, Harvard University, Cambridge, Massachusetts

Manabu Tanifuji Brain Science Institute, RIKEN, Wako, Saitama, Japan

Kazushige Tsunoda Brain Science Institute, RIKEN, Wako, Saitama, Japan

David Van Essen Department of Anatomy and Neurobiology, Washington University School of Medicine, St Louis, Missouri

Claire Wardak Institut des Sciences Cognitives, Bron, France

Dengke Xiao School of Psychology, University of St. Andrews, Fife, Scotland, United Kingdom

Yukako Yamane Brain Science Institute, RIKEN, Wako, Saitama, Japan

Karl Zilles Institute of Medicine, Research Center Jülich, Germany and C. & O. Vogt Brain Research Institute, University of Düsseldorf, Germany

Index