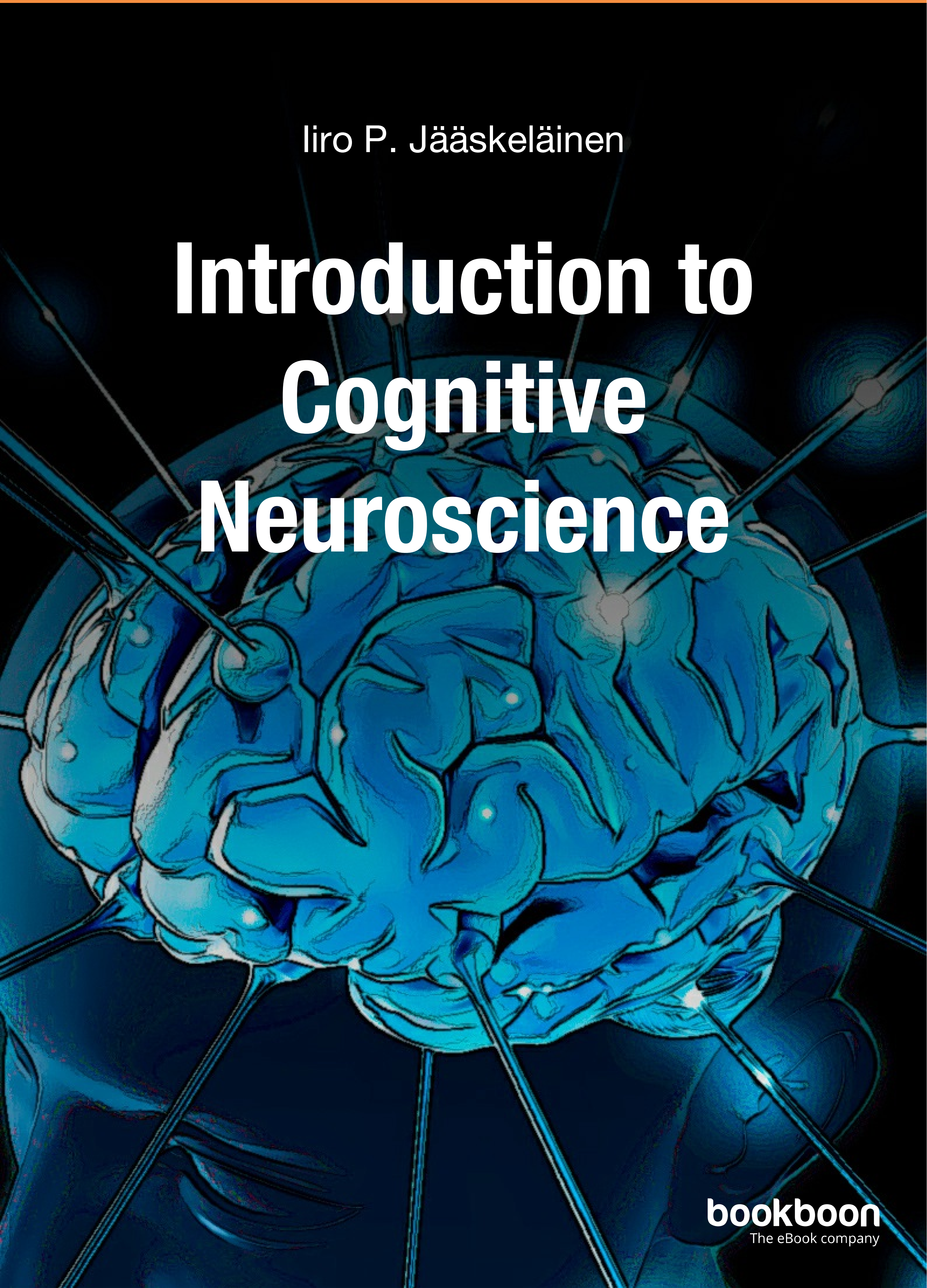
Iiro P. Jääskeläinen

Introduction to Cognitive Neuroscience



bookboon
The eBook company

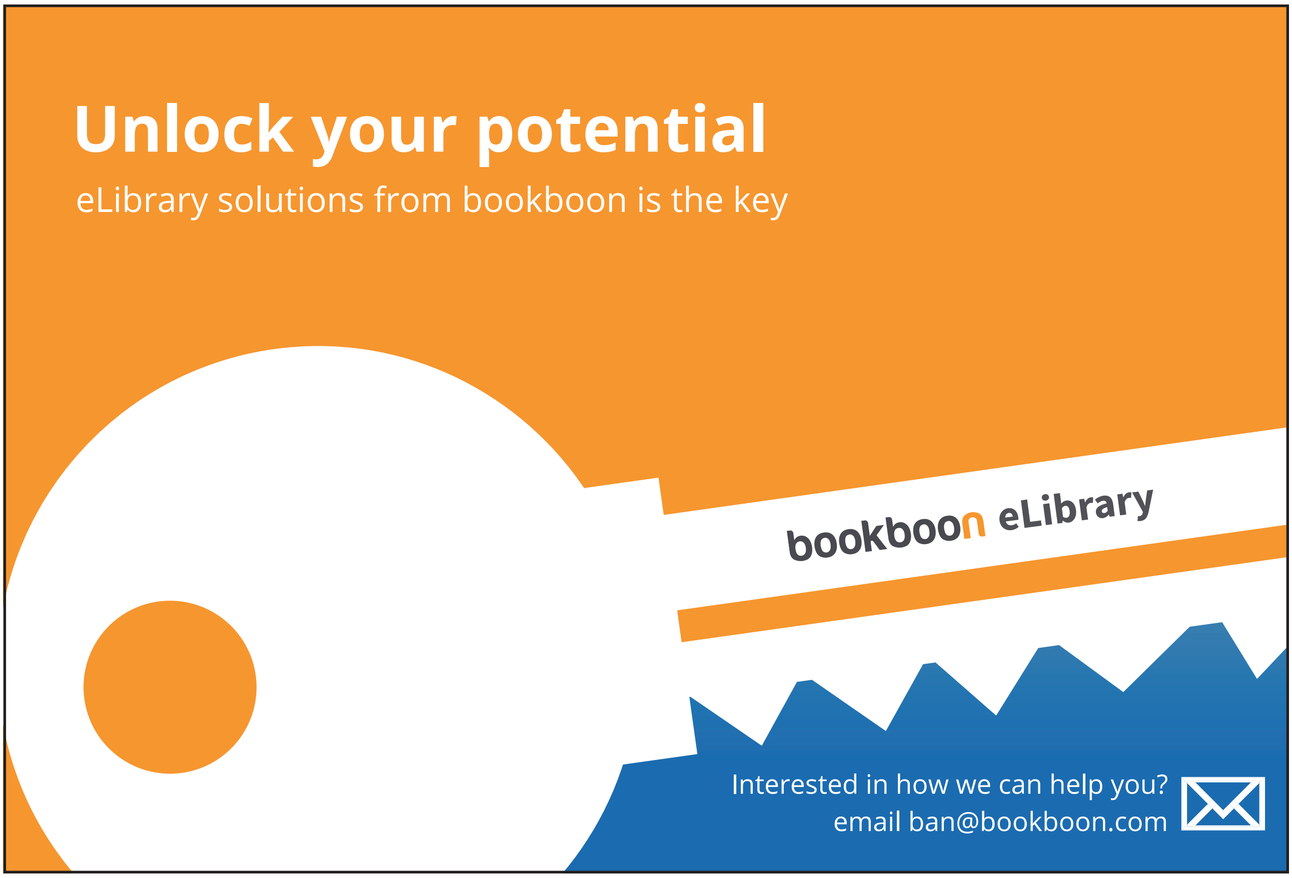
Iiro P. Jääskeläinen

Introduction to Cognitive Neuroscience

Introduction to Cognitive Neuroscience

ISBN 978-87-403-0100-7

Contents



Unlock your potential
eLibrary solutions from bookboon is the key
bookboon eLibrary
Interested in how we can help you?
email ban@bookboon.com

About the Author

Iiro P. Jääskeläinen obtained his Ph.D. in 1995 at the Department of Psychology, University of Helsinki, Finland. Currently, he is a cognitive neuroscientist at the Brain and Mind Laboratory at the Department of Biomedical Engineering and Computational Science in Aalto University School of Science, Espoo, Finland. Prior to joining Aalto University, Dr. Jääskeläinen worked as a post-doctoral fellow at the University of Helsinki and as an Instructor in Radiology at Harvard Medical School and Massachusetts General Hospital, Boston, USA.

Jääskeläinen has conducted cognitive neuroscience studies since the mid-1990s using EEG, MEG, fMRI, and psychophysics, with a relatively wide range of interests from perceptual and cognitive functions, including emotions and social cognition, in healthy humans to cognitive deficits in psychiatric and neurological disorders. He has also been involved in psychopharmacological studies addressing the neurotransmitter basis of human perceptual and cognitive functions.

Jääskeläinen has authored over 80 articles in international peer-reviewed journals including PNAS, Trends in Neurosciences, Biological Psychiatry, Journal of Neuroscience, Neuroimage, Human Brain Mapping, PLoS ONE, and Neuropsychopharmacology. He is also a review editor of Frontiers in Auditory Cognitive Neuroscience and an editorial board member of the Open Journal of Neuroscience.

For any feedback/thoughts on the Introduction to Cognitive Neuroscience, the author can be directly contacted by sending email to iiro.jaaskelainen@aalto.fi

1 What is cognitive neuroscience?

Cognitive neuroscience is a fairly young but rapidly growing discipline of science that aims at solving the intriguing question of how the brain (or, more generally, the body) gives rise to the human mind; how the brain enables one to think, plan, remember, understand other people, see, hear, and move about. Solving this so-called mind-body problem has been a central quest ever since the days of ancient Greek philosophers such as Plato, though many of the early philosophers mistook the brain as an organ merely responsible for cooling the body and not as the seat of thoughts, memories, emotions, and motivation that are collectively nowadays referred to as cognition. It is today, with the aid of modern neuroimaging technology, that it is becoming increasingly possible to shed light on the neural basis of cognitive functions.

To begin the exciting adventure into the depths of brain-mind interrelationships, it is useful to first imagine some events that could take place during an ordinary day. In the following, an example is provided taken from the life of a fictional university student. When reading through the example below, try to imagine what type of mental processes the student has to utilize while going through various events of the day. And once you have read through the whole book, you might wish to return to this short description of the day of a student, and figure out how the different events relate to the various cognitive functions that are described in the various chapters of this book.

Cognitive neuroscience in everyday life: an example from a student's day

Imagine a student, a responsible and well-mannered young adult named James, waking up on the morning of an important examination. James gets up, and while still mentally reciting the most central topics and details for the examination he goes through preparing breakfast and commuting to his college. On his way there, he receives a call from his loved one, who says that there is a need to seriously talk about something in the evening. That adds to the anxiety level of poor James and he starts going through in his mind different scenarios of what it might be that is on his/her loved one's mind. When doing so, he recollects interactions with his loved one that have taken place in the recent past, looking for memories of important cues; has there been signs of boredom, annoyance, or is she perhaps willing to take the next step in their relationship?

At the steps of the college, James bumps into an old friend from elementary school and immediately remembers what type of personality his friend has and all those things he did together with him. But James remembers that he has to hurry and thus cannot linger in a conversation. Moving on, James finally makes it to the hall where the examination is to take place, barely in time. As an unpleasant surprise, James notices that there is no one there! He frantically starts searching for any clues and finds a notice stating that the place of the examination has changed to another hall with the number of the hall indicated. With effort James recalls where this other hall is located within the confines of the campus and mentally calculates the fastest route to get there. James decides to run as there is only few minutes left and manages to make it in time.

Once the examination begins, James feels that he is in too much adrenalin to answer the examination questions, however, after a few minutes he manages to calm down and focus on the questions. Once the examination is over James is on elevated moods as he feels having done very well. At the same while as James is feeling rewarded that his hard studying has paid off, the memory of the morning phone call from his loved one creeps back into his mind, and James starts worrying about what is going to happen in the evening, what his loved one possibly wants to discuss about.

Mental phenomena that take place in the types of everyday life situations and tasks that James encounters and engages in (*e.g.,* planning one's actions, figuring out the intentions of other persons, maintaining one's goals persistently, feeling emotions, and memorizing things) are the very focus of cognitive neuroscience research. Each of these cognitive functions and the underlying neural mechanisms will be described in detail later in this book. In order to begin from the very basics, however, a more thorough definition of the word "cognitive" will be provided in the following.

Cognition is a general term encompassing all mental functions

Cognition is a rather general term that refers to all mental processes, such as perception, thinking, memory, motivation, attention, emotions, ability to understand the intentions and thoughts of other people, decision-making and self-awareness – in fact cognition also refers to the ability of one to read and comprehend this text, as well as the abilities that have been used to put this text together. In cognitive psychology, which is a somewhat older discipline than cognitive neuroscience, these mental processes have been studied without necessarily considering the possible underlying neural mechanisms.

In cognitive neuroscience, the goal is to elucidate the neural basis of the cognitive processes and, naturally, understanding the underlying neural mechanisms also helps one better understand the nature of the cognitive processes themselves. The relatively young discipline of cognitive neuroscience has emerged especially due to rapid developments in non-invasive neuroimaging methods over the last few decades (see Chapter 2) that allow quantification of brain activity in healthy volunteers while they are engaged in various perceptual and cognitive tasks. But one might wonder why it is important to know about cognitive functions and the underlying neural mechanisms? This important question is answered in the following.

1.1 Why is cognitive neuroscience important?

Solving the mind-body problem has been a fundamentally intriguing question since the days of ancient Greek philosophers and indeed research on the neural basis of the human mind is clearly warranted as a pure basic research venture, however, cognitive neuroscience also yields manifold benefits to applied research and other fields of science. For example, attempts to build artificial intelligence and robots have already benefited from the knowledge of cognitive processes and the underlying neural mechanisms. In many ways the brain is a highly efficient computer that is able to carry out many information-processing tasks more quickly than the most powerful of man-made computers. Thus, gaining deeper understanding of the core processing principles and organization of brain functions will offer important insights to computer and computational sciences.

Clinical research into disorders of the brain where cognitive processes are deficient, such as various types of dementias, severe mental health disorders such as schizophrenia, developmental disorders such as autism, and problems experienced by people with brain damage, are also benefited by knowledge acquired in cognitive neuroscience basic research. Clinical research often aims at development of drugs that alleviate cognitive deficits that patients suffer from and thus the fruits of cognitive neuroscience research are also widely utilized by pharmaceutical companies in their vast research and development efforts.

Development of various types of man-machine interfaces in engineering is another rapidly growing application area where cognitive neuroscience plays a central role. This includes design of human-computer as well as human-cell phone interfaces. The latter includes, for example, design of touch screens, vibration alarms, and various attempts towards offering enhanced reality, for instance, to enhance the awareness of the user to attractions such as restaurants of preferred ethnicity that are nearby. The first brain-computer interfaces have also been recently implemented by computer game industry where the player partly directs the actions in the game with his/her brain activity that is measured with specific sensors placed on the scalp of the player.

1.2 The brief history of cognitive neuroscience

How did the discipline of cognitive neuroscience emerge and develop? In the following, the most important historical landmarks will be briefly described that led into the development of this exciting scientific discipline. It is important to note that while research on the interrelationships between the brain and mind have not been called cognitive neuroscience until relatively recently, the first attempts to map correspondence between the brain and the mind took place already hundreds of years ago. Sometimes even the early attempts of the Greek philosopher Aristotle to connect mind and the heart are mentioned as initial steps towards localization of mental functions, even though it was Plato, the mentor of Aristotle, who believed the brain to be the seat of mental processes.

In the 17th Century Europe there were theoretical formulations advanced by Rene Descartes of nerves containing animal fluids, the movement of which were thought to carry sensory and motor information, however, the realization that neurons (*i.e.*, specialized cells that make up the nervous system) are important for cognition (the “neuron doctrine”) emerged thanks to staining methods developed by Italian scientist Camillo Golgi. These methods allowed one to visualize neurons in tissue samples (Golgi, 1873, Pannese, 1996), for an example of this, see **Figure 1-1** below. These staining methods were utilized by Ramón y Cajal in his subsequent work (Ramón y Cajal, 1899, 1904, Andres-Barquin, 2001). Together, Golgi and Ramón y Cajal received the Nobel Prize for their work in 1909. In the neuron doctrine, which still is valid today, it is assumed that a single neuron is the elementary building block of cognition (neurons are described in detail in Chapter 4). Despite of these significant advances, it was not until the 19th Century that research specifically addressing brain-mind interrelationships begun emerging.

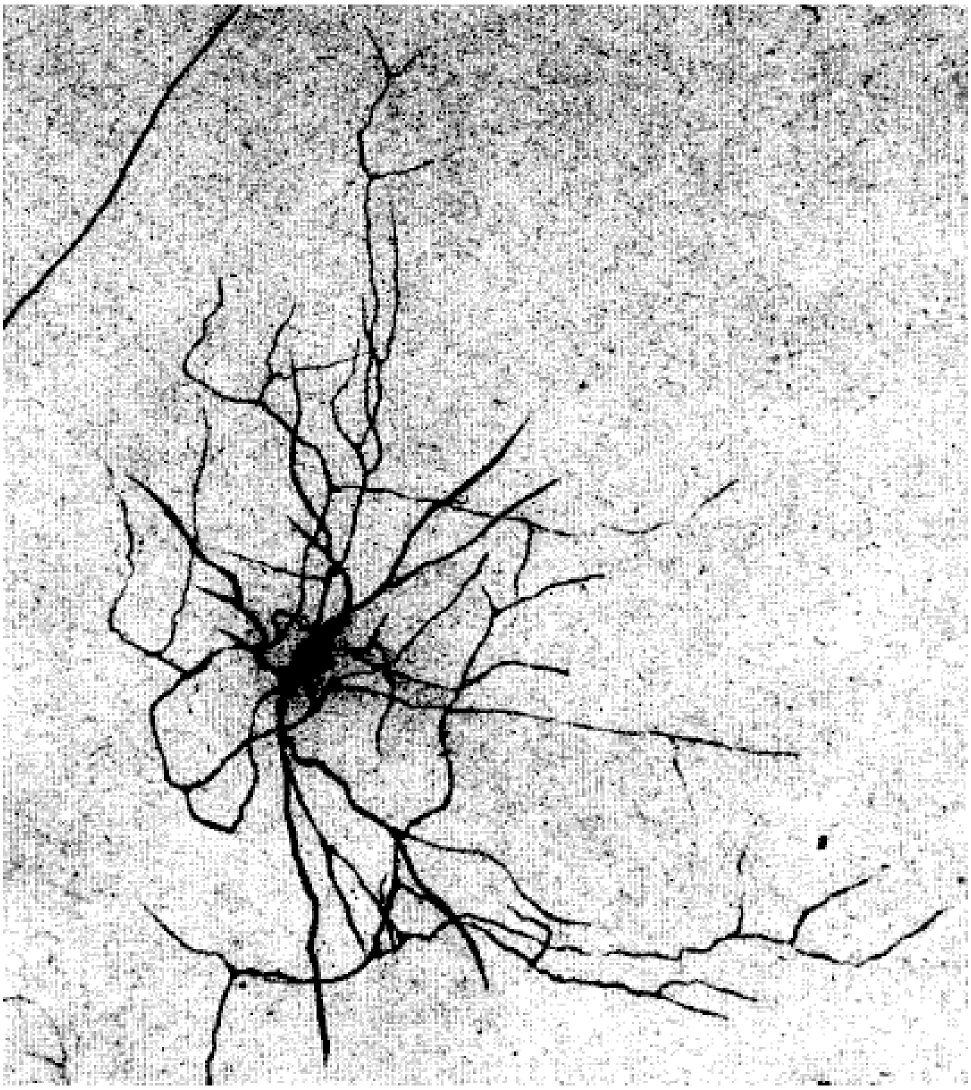


Figure 1-1. A multipolar neuron from the spinal cord of mammalian fetus impregnated using Golgi's "black reaction" staining method (adapted from (Pannese, 1996)).

Phrenology: an early attempt to localize mental functions

The first attempt to specifically connect specific brain regions with mental functions was introduced by German scientists Franz Joseph Gall and Johann Gaspar Spurzheim, who started the discipline of phrenology in the turn of the 18th and 19th Centuries (Gall and Spurzheim, 1809, Simpson, 2005). Phrenology was based on mapping the interrelationships between depressions and bumps in the skull with various personality features (see **Figure 1-2**). The idea behind this was that more developed brain areas, which would go with more developed aspects of personality, would create small bumps in the skull. Thus, it was surmised that by studying the correlations between personality features (such as "hopefulness" and "firmness") and the bumps in the skull one would be able to construct a brain map of personality. While phrenology has been today most often mentioned as a good example of pseudoscience, it is one of the first documented instances where there was an attempt to localize aspects of mental processes to the brain.

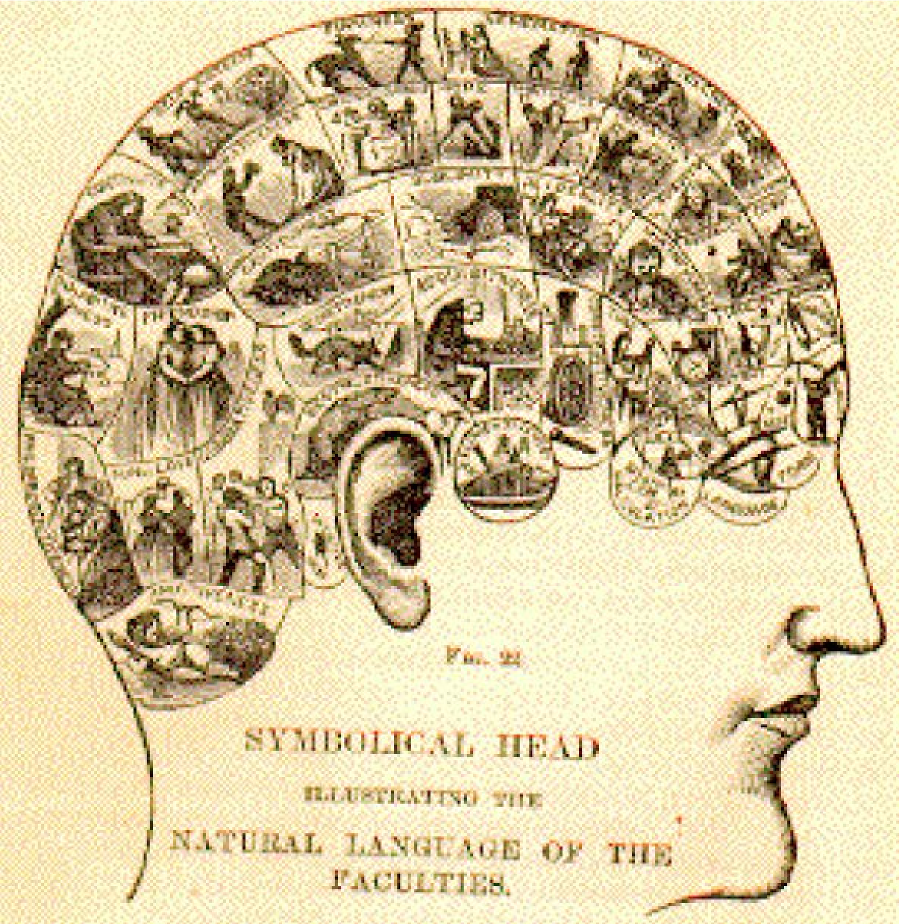


Figure 1-2. In phrenology that flourished in the turn of the 18th and 19th Centuries it was assumed that small bumps in the skull reflected more developed brain areas underneath and that finding associations between such bumps and personality features would be possible, thus revealing aspects of localized brain functions. While phrenology is nowadays regarded as a pseudoscience, the basic idea that cognitive functions can be localized in the brain has become one of the guiding principles in modern cognitive neuroscience (adapted from http://www.cerebromente.org.br/n01/frenolog/frenologia.htm).

Early animal lesion studies challenged the localized-functions hypothesis

The basic assumptions of phrenology were soon challenged. Pierre Flourens carried out series of experiments in Paris where he observed the effects of localized brain lesions on the behavior of rabbits and pigeons. While he did see some marked differences in cases where, for instance, the whole cortex or cerebellum were removed, he could not find strict relationships between various parts of the cortex and specific mental functions as suggested by the phrenologists (Flourens, 1824). Based on his work, Flourens proposed that the brain is an aggregate field, meaning that cognition and memory are represented in the brain in a distributed manner. The aggregate field view that was put forth by Pierre Flourens was the first instance of holism standpoint in a theoretical debate between localized *vs*. holistic functions.

The localization *vs*. holism debate has in fact been one of the major theoretical disputes in cognitive neuroscience. A century after Flourens, there was prominent work by Karl Lashley who carried out lesion studies in animal models and strongly argued against the strict localizationist views of brain function (Lashley, 1931). Today, it has become obvious that both sides have been partly correct in their views; while it is widely recognized that there is a high degree of specialization between different cortical areas, it is also true that cognitive functions arise due to joint activity of a network of brain areas, and that a given cortical area can be recruited in different types of cognitive tasks, depending on the specific processing requirements of the task. As an example of this, the parts of cortex that process sensory information are also part of the network of brain areas underlying one's ability to imagine things, for instance, the visual cortex that underlies the ability to see is also activated during mental visual imagery (Kosslyn et al., 1993).

Brain lesions and language impairments: Broca and Wernicke

Paul Broca and Carl Wernicke documented the very first observations of specific cognitive impairments following focal brain damage in the 19th Century. Broca presented findings of a patient who had lost his ability to speak (this patient could only utter a single word "tan"), yet the patient could still understand what was spoken to him, suggesting that there was a specific loss of speech production ability (Broca, 1861). Broca followed up the patient and conducted an autopsy when the patient later deceased. He then observed that a portion of the left inferior frontal lobe was lesioned in this patient (see **Figure 1-3**). Today the area of inferior frontal lobe is still widely referred to as Broca's area. Soon after the findings described by Paul Broca, another scientist Carl Wernicke documented a patient who had lost his ability to understand speech (Wernicke, 1874). Autopsy revealed a lesion in the posterior aspects of the patient's temporal lobe. This area is still today referred to as the Wernicke's area.

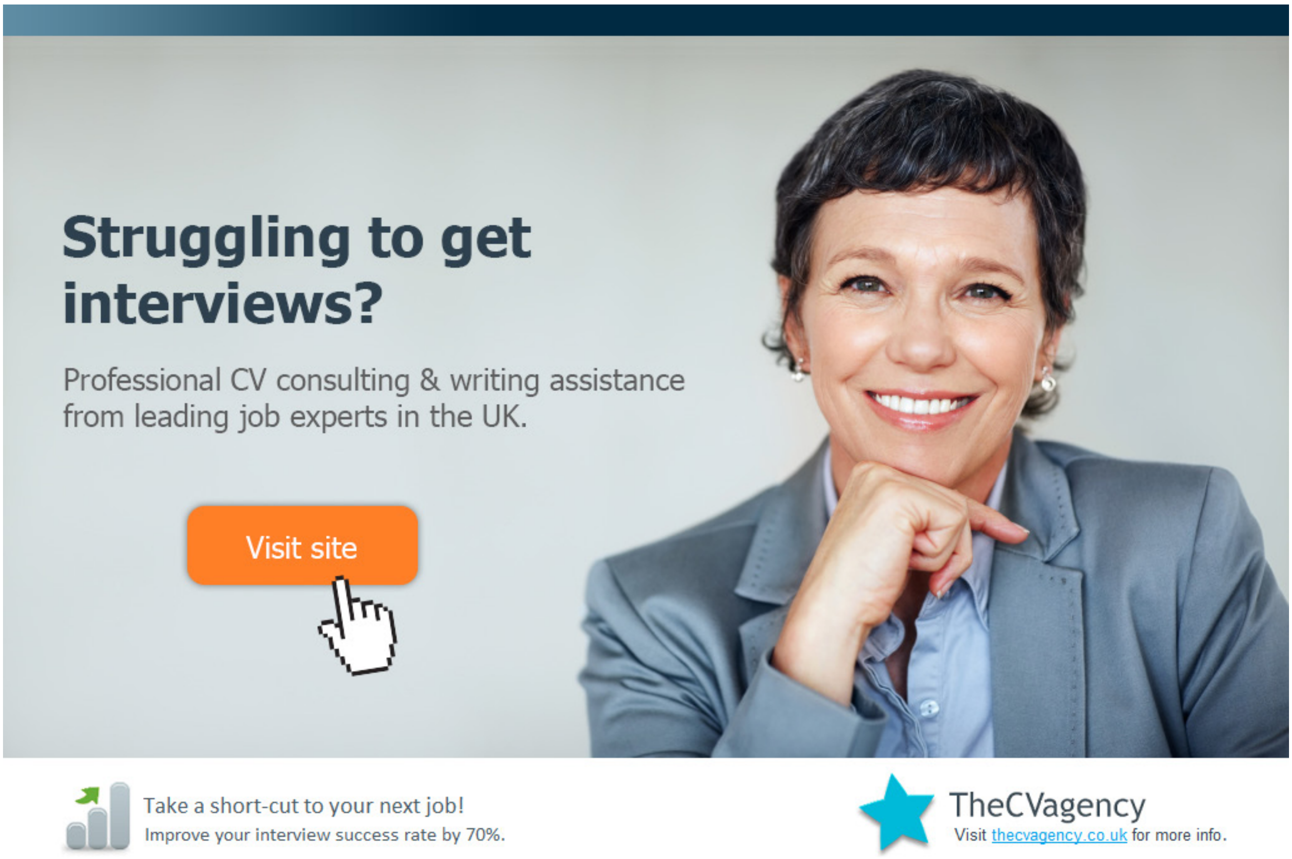


Figure 1-3. Shown is lateral view of the left side of the brain of Paul Broca's patient with the lesion in the left inferior frontal cortex encircled (adapted from http://neurophilosophy.wordpress.com/2007/04/26/old-brains-new-ideas/).

Together, these findings suggested distinct localized brain regions that support speech production and comprehension. In the late 19th Century, these findings moved the scientific debate back to the direction of localized brain functions, and this time the methods, unlike in the case of phrenology, were scientifically valid. Naturally, in the early days of Broca and Wernicke the mapping of correspondence between locations of brain lesions and specific cognitive impairments was tedious work, since there were no modern neuroimaging methods such as magnetic resonance imaging (that is described in Chapter 2) to map the lesion sites non-invasively but rather scientists had to wait until a patient died to conduct autopsy. This not only slowed research down but also reduced the number of cases reported and thereby raised concerns of whether the results could be generalized. Nonetheless, through the pioneering observations of Paul Broca and Carl Wernicke in the 19th Century, a tradition of documenting interrelationships between brain lesion sites and cognitive deficits emerged.

Behavioral tests of cognition

Importantly, in order to study cognitive deficits in neurological patients, behavioral tests had to be developed that measure specific cognitive functions. In other words, one could not rely on vague descriptions such as "the memory of the patient seemed to be compromised", but rather specific tests that measure memory functions quantitatively needed to be developed in order to be able to assess to what extent (and which) memory functions were impaired. The development of such behavioral tests (also known as neuropsychological tests and intelligence tests), and development of behavioral paradigms to assess perceptual and cognitive functions ("psychophysics"), has been vital for the development of modern understanding of cognitive functions. In fact, until the development of functional imaging methods in the late 20th Century, information gathered in studies using behavioral tests in healthy volunteers and in neurological patients made up the bulk of what was known about cognitive functions and the underlying neural mechanisms.

Development of non-invasive neuroimaging methods

Development of non-invasive functional neuroimaging methods, starting with electroencephalography in the early 20th Century and then, towards the end of the 20th Century, rapidly expanding repertoire consisting of positron emission tomography, magnetoencephalography, functional magnetic resonance imaging, near-infrared spectroscopy, and transcranial magnetic stimulation, has led to the emergence of cognitive neuroscience as a rapidly expanding discipline (these methods are described in Chapter 2). Recently, development of mathematical methods with which it is possible to combine complementary information provided by different neuroimaging methods has proven highly useful (Liu et al., 1998, Auranen et al., 2009).

Indeed, the ability to peek inside the brain of healthy subjects and track electromagnetic and metabolic activity within discrete anatomical structures with a high temporal resolution during performance of perceptual and cognitive tasks is truly amazing. However, it can be said that the vast potential offered by these methods is still far from having been fully exploited. As is always the case with novel research methods, it takes time before the various ways that the methods can be used are realized. The signal-to-noise ratio of neuroimaging equipment is also rapidly improving through technological advances.

On the other hand, constant development of analysis algorithms makes it possible to design novel experimental setups. As an example of this, it was not until 2004 that it was discovered that recording highly meaningful across-subjects replicable brain responses from sensory receptive cortical areas of the brain is possible when experimental subjects are freely viewing a highly naturalistic stimulus such as a feature film (Hasson et al., 2004) and subsequently it was shown that inspection of across-subjects correlated brain activity under such conditions could be also seen in the prefrontal cortical areas (Jaaskelainen et al., 2008) that support higher cognitive functions, including how one makes decisions and how one manages to infer intentions of other people.

1.3 The present and exciting future of cognitive neuroscience

It is often said that nothing is more difficult than predicting the future, however, it is tempting to speculate, based on a number of highly exciting recent developments, where cognitive neuroscience might be moving in the near future. Constant development of more sensitive and accurate neuroimaging and data analysis methods creates new research possibilities. Advances in theoretical and computational models is, however, equally important as such models allow one to derive testable hypotheses and guides the analysis of vast datasets that are nowadays obtained in neuroimaging experiments. With the help of advances in both methods and theory, cognitive neuroscientists are constantly devising novel experimental paradigms that allow investigation of phenomena that could not have been effectively investigated before.

Mind-reading using neuroimaging methods

The 21st Century has seen the emergence of, at this stage rather crude but nonetheless promising, mind-reading methods. As an example of this, scientists managed to show recently how movie clips that healthy volunteers were watching during functional magnetic resonance imaging could be reconstructed based on the brain activity that was recorded (Nishimoto et al., 2011). Examples of such reconstructed clips are shown in **Figure 1-4** below. Typically the "mind-reading" studies have used sophisticated machine learning algorithms where distributed patterns of brain activity are associated with specific mental states and thoughts of the subjects.

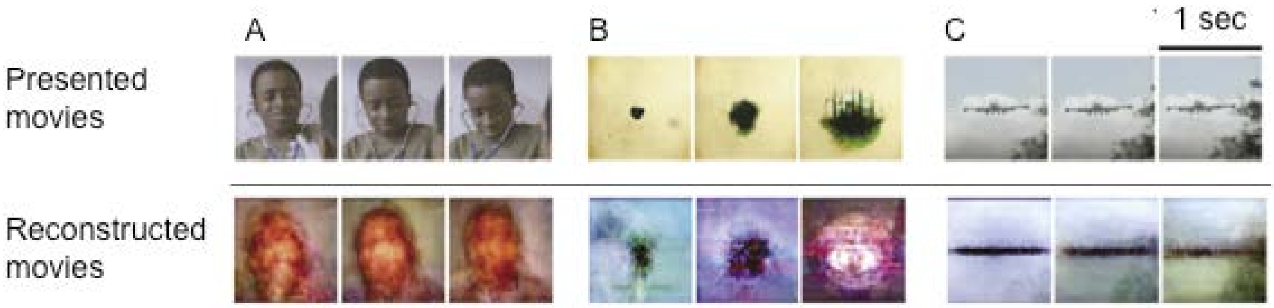


Figure 1-4. Reconstruction of movie clips from measured brain activity. On top are shown three exemplary movie clips presented to subjects, and at the bottom, movie clips reconstructed based on measured brain activity (adapted from (Nishimoto et al., 2011)).

Again, the fact that distributed brain activity, rather than activity of any single brain area, seem to underlie mental states and cognitive functions provides another indication of how the localized functions and holistic hypotheses are both partially correct. It is networks of brain areas, and networks of neurons within each of the networked brain areas, that give rise to cognition. On the other hand, given that it is possible to quantify these brain network states, a more fundamental understanding of cognition and the underlying neural mechanisms can be achieved. This will also provide unprecedented possibilities for various applied research areas of cognitive neuroscience such as investigation of deficits in brain function in various dementias or psychiatric disorders.

Brain-computer interfaces constitute one example of an application area that has already been made possible with mind-

reading neuroimaging techniques (Brunner et al., 2011). In brain computer interfaces, computational algorithms learn specific patterns of brain activity corresponding to a few mental commands that subjects are making. Once learning has been accomplished, it is possible to guide the computer interface with the mental commands. The most important application area of brain computer interfaces is helping paralyzed patients who have lost their ability to communicate. While existing brain computer interfaces cannot make smooth and lively communication possible, any communication ability is helpful for such patients.

Research on the neural basis of emotions and higher cognitive functions

Another rapidly emerging aspect of modern cognitive neuroscience is that developments in neuroimaging and analysis methods are making it possible to use highly naturalistic stimuli such as movies and computer games instead of more traditional highly controlled experimental setups involving repetitive showing of the same, often highly simplified, stimulus to the subjects (*i.e.*, repeating the same stimulus / task has been necessary in order to attain sufficient signal-to-noise ratio and, on the other hand, keeping the stimulus simple enough has been necessary to keep the neural responses from being too ambiguous to analyze with the data analysis approaches that have been available). This is making it possible to study higher cognitive functions such as goal-directed behavior, decision-making and emotions under more ecologically valid conditions.

1.4 The purpose of this book

The purpose of this book is to provide a comprehensive introduction to the basic concepts of cognitive neuroscience. While the book is primarily intended for university students in various disciplines (*e.g.*, psychology, medicine, and engineering) taking their first introductory course in cognitive neuroscience, it also constitutes a source of introductory-level information for any person who is interested in the study of mind-brain relationships. Of course, in order to understand the cognitive functions and the underlying neural mechanisms, and the roles they play in everyday life, it is necessary to begin from the basic concepts. Since cognitive neuroscience is to a large degree about studying where different functions are located in the brain and how different brain areas carry out cognitive tasks together, knowing the names of various anatomical structures makes it possible to understand the findings that are described in subsequent chapters of this book. In the next chapter, the most central anatomical structures of the brain will be described.

1.5 References

ANDRES-BARQUIN, P. J. 2001. Ramón y Cajal: a century after the publication of his masterpiece. *Endeavour,* 25**,** 13-17.

AURANEN, T., NUMMENMAA, A., VANNI, S., VEHTARI, A., HÄMÄLÄINEN, M., LAMPINEN, J. & JÄÄSKELÄINEN, I. P. 2009. Automatic fMRI-guided MEG multidipole localization for visual responses. *Hum Brain Mapp,* 30**,** 1087-1099.

BROCA, P. 1861. Remarque sur le siege de la faculté du language articulé, suivie d'une observation d'aphémie (perte de la parole). *Bulletin de la société anatomique de Paris,* 36**,** 330-356.

BRUNNER, P., BIANCHI, L., GUGER, C., CINCOTTI, F. & SCHALK, G. 2011. Current trends in hardware and software for brain-computer interfaces (BCIs). *Journal of Neural Engineering,* 8**,** 025001.

FLOURENS, P. 1824. *Recherches experimentales sur les propriétés et les fonctions du système nerveux dons les animaux vertébrés,* Paris, Chez Crevot.

GALL, F. J. & SPURZHEIM, G. 1809. *Anatomie et physiologie système nerveux en general et du cerveau en particulier,* Paris, Schoell.

GOLGI, C. 1873. Sulla struttura della sostanza grigia del cervello. *Gazzetta Medica Italiana-Lombardia,* 6**,** 244-246.

HASSON, U., NIR, Y., LEVY, I., FUHRMANN, G. & MALACH, R. 2004. Intersubject synchronization of cortical activity during natural vision. *Science,* 303**,** 1634-1640.

JAASKELAINEN, I. P., KOSKENTALO, K., BALK, M. H., AUTTI, T., KAURAMÄKI, J., POMREN, C. & SAMS, M. 2008. Inter-subject synchronization of prefrontal cortex hemodynamic activity during natural viewing. *The Open Neuroimaging Journal,* 2**,** 14-19.

KOSSLYN, S. M., ALPERT, N. M., THOMPSON, W. L., MALJKOVIC, V., WEISE, S. B., CHABRIS, C. F., HAMILTON, S. E., RAUCH, S. L. & BUONANNO, F. S. 1993. Visual mental imagery activates topographically organized visual cortex: PET investigations. 5**,** 263-287.

LASHLEY, K. 1931. Mass action in cerebral function. *Science,* 73**,** 245-253.

LIU, A. K., BELLIVEAU, J. W. & DALE, A. M. 1998. Spatiotemporal imaging of the human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations. *Proc Natl Acad Sci USA,* 95**,** 8945-8950.

NISHIMOTO, S., VU, A. T., NASELARIS, T., BENJAMINI, Y., YU, J. & GALLANT, J. L. 2011. Reconstructing visual experiences from brain activity evoked by natural movies. *Current Biology,* 21**,** 1641-1646.

PANNESE, E. 1996. The black reaction. *Brain Research Bulletin,* 41**,** 343-349.

RAMÓN Y CAJAL, S. 1899, 1904. *Textura del Sistema Nerviosa del Hombre y de los Vetebrados,* Madrid, Spain, Moya.

SIMPSON, D. 2005. Phrenology and the neurosciences: contributions of F.J. Gall and J.G. Suprzheim. *ANZ J Surg,* 75**,** 475-482.

WERNICKE, C. 1874. *Der aphasische symptomencomplex: eine psychologische studie auf anatomischer basis,* Breslau, Cohn & Weigert.

2 Methods of Cognitive Neuroscience

Given the vast complexity of human perceptual and cognitive processes on one hand, and that of the human brain on the other (with approximately 10^{11} neurons with each neuron having on the average ~7000 synaptic connections with other neurons), it is clear that developments in research methodology play a pivotal role in the progress of cognitive neuroscience towards it's ultimate goal of elucidating the neural basis of perceptual, cognitive and emotional functions in both health and disease. Fortunately, progress in non-invasive functional neuroimaging methods that allow one to record brain activity during presentation of various stimuli and execution of tasks in healthy volunteers has been very rapid. Today, modern imaging technology, along with sophisticated signal analysis algorithms, provides possibilities for scientists that one could only have dreamt of a few decades ago, such as imaging patterns of brain activity that can be associated with specific thoughts.

The extent that the neuroimaging methods can be utilized depends largely on theoretical advances in cognitive neuroscience. This is because theories and models of the neural basis of perceptual and cognitive functions makes it possible for the cognitive neuroscientists to design experiments that test specific hypotheses. As an example of hypothesis testing, recall the early experiments conducted by Pierre Flourens described in Chapter 1. He set forth to critically examine hypotheses that he could formulate based on the theory of phrenology (*i.e.*, that various mental functions are strictly localized in the brain). Since he did not find any apparent relationships between site of lesion and behavioral deficits, his results argued against phrenology and he proposed the aggregate field model of brain function.

As another simple example of a research hypothesis one could assume (*i.e.*, hypothesize) that there are differences in height between men and women with men being taller. Then, by acquiring a large enough sample of male and female volunteers and measuring their heights, one would be able to test this hypothesis by comparing height distributions between the male and female samples. There are simple statistical tests that enable one to estimate probability at which the two samples differ. Typically in science the probability that two samples are similar has to be less than 0.05 (*i.e.*, less than a five percent chance) before it can be concluded that the samples significantly differ from one another.

Hypotheses are very important in cognitive neuroscience. Without a solid theoretical framework and specific hypotheses, recording of brain activity would be more or less random and the results that are obtained would be difficult to interpret meaningfully (save for specific data mining approaches that find consistencies across huge datasets – but even in the case of such studies, interpretation of the emerging patterns of analysis results requires theoretical knowledge of cognitive processes and how the human brain functions). It is also important to note that given the complexity of data obtained using the modern neuroimaging methods, it is quite typical that there are several alternative explanations to a given set of findings. One can then test between the alternative explanations by formulating specific hypotheses and running further experiments. The outcome of the hypothesis testing then advances the theory/model either by validating or disproving it. In the latter case, one has to amend the theory/model to account for the discrepancy. This is the way that science progresses.

Importantly, none of the cognitive neuroscience methods that are available today can provide all the necessary information to solve how the brain works. Rather, each of the methods should be seen as providing complementary information that helps solving the big puzzle. One highly useful viewpoint to the complementary nature of the methods that are used in the study of brain function is presented in **Figure 2-1** below. As can be seen, each of the methods is alone either temporally or spatially limited, but when information across the methods is combined, a relatively high spatiotemporal accuracy is achieved.

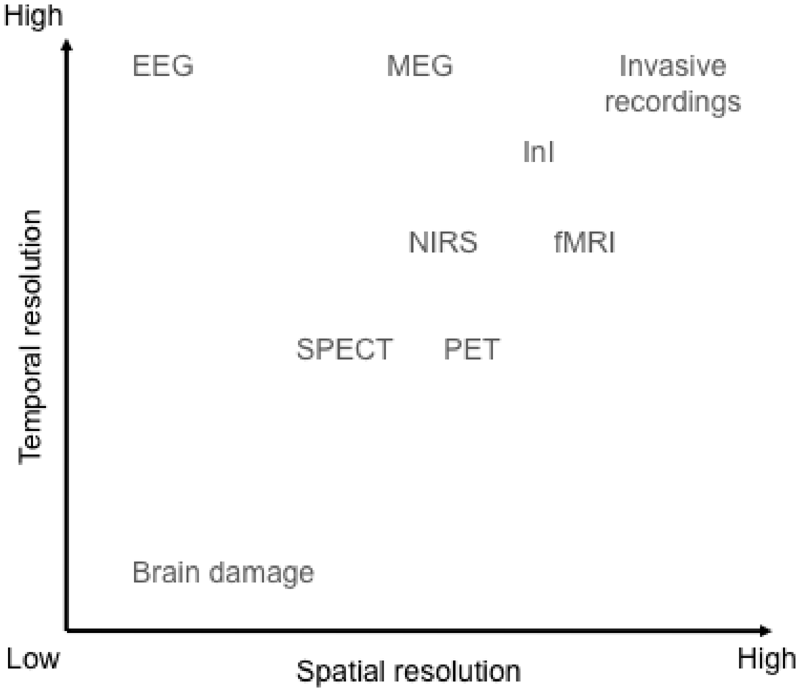


Figure 2-1. Methods that are used in cognitive neuroscience differ in terms of temporal and spatial resolution. Thus, combination of information provided by the different methods is an important aspect of cognitive neuroscience research. Note that invasive recordings are possible in human subjects only in relatively rare instances where treatment of patients suffering from severe medical conditions such as medication-resistant epilepsy necessitate undergoing surgery that requires invasive recordings. Abbreviations from top left: EEG = electoencephalography (measurement of scalp electric potential differences due to neuronal population activity), MEG = magnetoencephalography (measurement of weak extracranial magnetic field changes produced by electric activity of large populations of neurons), InI = inverse imaging (a novel method that speeds up conventional functional magnetic resonance imaging that measure blood flow and oxygen level changes that are coupled with local neuronal activity (Lin et al., 2006)), NIRS = near infrared spectroscopy (measurement of changes in optical properties of blood with changes in blood oxygen content that are coupled with neuronal activity), fMRI = functional magnetic resonance imaging (measurement of blood flow and oxygenation level based on tiny magnetic field changes produced by differential magnetic properties of oxygenated *vs*. deoxygenated red blood cells), SPECT = single positron emission computed tomography (measurement of weak radiation caused by a marker substance that is injected to blood flow of experimental subjects/patients), PET = positron emission tomography (measurement of weak radiation caused by marker substance that is injected to blood flow of experimental subjects/patients).

An analogy of a group of blind men inspecting an elephant is quite often used to describe the need for integration of information provided by different methods: each of the men, inspecting different parts of the elephant, one the trunk, one body, yet another the tail or ears, arrive alone at quite different types of interpretations of what type of an animal the elephant is. Similarly, by comparing results acquired with different methods, one can get a more complete picture of how the brain works. Naturally, knowing what each method is capable of, what they exactly measure, and what their limitations are is the key to being able to successfully integrate information across studies that use different methods.

Furthermore, neuroimaging methods differ in what specific aspect of brain function they measure. For example, some of the methods measure electric potentials and/or magnetic fields caused by large neuronal populations and some others measure the blood flow changes that are tightly coupled with local changes in neuronal activity. In the following, the most important methods of cognitive neuroscience will be described, beginning with specific behavioral measures that are obtained in healthy and brain damaged subjects. Notably, prior to invention of modern neuroimaging methods, behavioral studies have produced most of what is known about cognitive functions and their underlying neural mechanisms.

2.1 Behavioral measures of cognitive and perceptual functions

Behavioral measures of cognitive and perceptual functions can be roughly divided into two classes. On one hand, deficits in perceptual and cognitive processes can be assessed in patients with focal brain damage to reveal function-structure interrelationships, as well as to delineate which perceptual and cognitive functions are separable from one another. For instance, it has been observed that perception of visual objects and their spatial positions can be separately damaged, suggesting that these two aspects of visual stimulation are processed by independent brain mechanisms (Newcombe and Russell, 1969).

On the other hand, various stimulus and task manipulations can be utilized to indirectly probe neural basis of cognitive functions. For instance, neurons tire when they repeatedly face the same stimulus or task. This phenomenon has been termed adaptation, and it can be utilized to test whether certain types of information processing are served by distinct neuronal populations by testing to what extent adaptation using a certain class of stimuli affects, for instance, the speed of processing or threshold of detecting the same *vs.* another class of stimuli (Sekuler et al., 1968). Based on the pattern of observed adaptation effects observed, it can be inferred that there are distinct neural populations that process the two types of stimuli.

Figure 2-2 below illustrates an example of the type of stimuli that have been used in behavioral adaptation studies. The stimuli that are shown here have been used in studies of adaptation effects to gender of facial stimuli. Using morphing

techniques, a continuum has been created between male and female faces, with the midpoint being an androgyne face. The subjects can then presented either a male or a female face for a sufficiently long time to induce adaptation, followed by presentation of the androgyne face and a request to determine whether the androgyne face is a male or a female. After viewing a male adaptor face, the subjects tend to perceive more often the androgyne face as a female and after viewing (*i.e.*, adaptation to) the female face, the androgyne face is more frequently perceived as a male. Findings from studies using this type of stimuli suggest that there are neurons specifically sensitive to gender-related facial features (Webster et al., 2004).

Figure 2-2. Shown is an example of stimuli that have been used in behavioral adaptation studies. Here, a morphed face (shown in the middle) that represents the midpoint between male and female faces is perceived more often as a female face after viewing the male adaptor face, and *vice versa*. This suggests that there are neurons coding for male/female gender cues in facial stimuli and that when the neurons sensitive to cues of one of the genders tire, response of the remaining neurons will dominate processing of the gender-ambiguous face. Adapted from (Rhodes et al., 2004).

Brain damage occurs in humans due to a variety of causes

Brain damage in humans can occur due to a variety of reasons. While brain damage due to head wounds caused by, for instance, bullets, are very common during wartime, strokes and brain tumors are by far the most common causes of brain damage during times of peace. Stroke can mean both an arrest of circulation due to a clot blocking a cerebral vessel that results in neuronal death due to anoxia, or anoxia and other complications due to rupture of a vessel that results in hemorrhage. Strokes differ significantly from brain tumors in that their onset is sudden, whereas in the case of tumors, the condition develops often quite slowly and typically symptoms do not appear until the tumor has grown to a relatively large one. Thus, surgical removal of the tumor typically involves removal of a relatively large part of healthy brain tissue around the tumor.

Medication-resistant epilepsy is another case where surgical removal of brain tissue is used. Head traumas and associated brain damage that occurs in, for instance, traffic accidents constitute another relatively common cause of brain damage. In such accidents, violent movements and twisting of the brain mass within the cranial cavity often result in tearing of neuronal connections between brain regions, as well as focal damage underneath the part of the skull that received the impact and on the opposite side of the brain due to the brain hitting/pressing against the opposite wall of the cranium during the impact. This latter effect is also known as "coup counter". Finally, there are a variety of neuropsychiatric and neurological disorders, such as attention-deficit hyperactivity disorder, schizophrenia, Parkinson and Alzheimer's diseases where brain function is abnormal.

Behavioral measures of cognitive deficits following brain damage

As was mentioned in Chapter 1, French physician Paul Broca was among the first to document a specific cognitive deficit following brain damage (Broca, 1861). Broca was presented with a patient who suddenly had lost his ability to speak. He could still understand verbal instructions but could only utter the word "tan". After the patient had died, Broca conducted an autopsy, documenting which parts of the brain had been damaged, and thus potentially of importance for speech production. What Broca observed was that an area of inferior frontal lobe of the left hemisphere had been destroyed in this patient. Today, Broca's aphasia (aphasia meaning inability to speak) is still a valid term that refers to difficulties in speech production. Pars opercularis and pars triangularis of the inferior frontal gyrus are collectively referred to as Broca's area. For an approximate location of the Broca's area, see **Figure 2-3** below; for an overall introduction to macroanatomy of the brain, see Chapter 3.

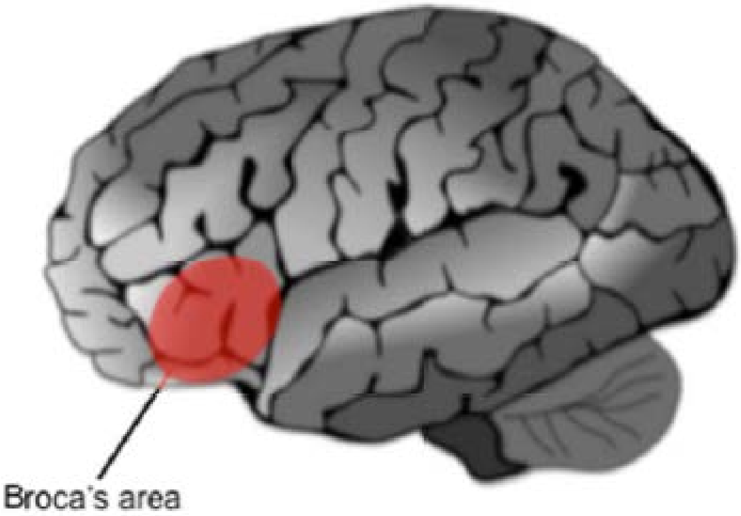


Figure 2-3: The approximate location of Broca's area, the damage of which was documented by Paul Broca to be associated with loss of ability to produce speech.

This early finding gave birth to an entire field of study where cognitive deficits in brain damaged patients are measured using various behavioral tests. An Austrian neurologist Carl Wernicke published, after the report of Broca, another case of aphasia, where the patient was no longer able to understand speech, and his speech output was incomprehensible (Wernicke, 1874). Autopsy revealed that an area located in the posterior-superior-lateral part of the temporal lobe was damaged in this patient. Wernicke called this deficit sensory aphasia, and today the condition is still widely referred to as Wernicke's aphasia.

Naturally, this research tradition, based on documenting cognitive deficits in patients with focal brain damage, has vastly progressed since the early days of Paul Broca and Carl Wernicke. Behavioral tests have been developed and refined into intricate psychophysics paradigms and neuropsychological test batteries (Lezak et al., 2004) that measure specific aspects of cognitive and perceptual functions such as memory, executive functions, and selective attention (see the respective Chapters of this book for more detailed description of the behavioral tasks that are used in these tests).

Development of structural imaging methods helped mapping lesion sites

Increasingly, patients with focal brain damage are also studied using modern neuroimaging techniques, which allows for investigation of the effects of focal brain lesions on brain function. Development of neuroimaging methods has helped also in localization of the lesion sites; whereas during the early days of Paul Broca and Carl Wernicke researchers had to wait for the patient to decease (which might take decades) in order to carry out autopsy (if they obtained permissions to do so) and see which areas of the brain had been damaged, modern neuroimaging methods allow one to non-invasively image the living brain in three dimensions and thus relatively effortlessly document the loci of brain damage.

Low number of patient cases and variable lesion sites reduces power of patient studies

There are inherent limitations to the approach where effects of brain damage on cognition are mapped to study the neural basis of cognitive functions. First of all, brain damage is rarely circumscribed, but rather involves extensive volumes containing multiple areas. Thus, it is difficult to associate a specific cognitive deficit with specific brain structures, especially as the number of cases that can be found for a given study is typically rather low. Given that there is inter-individual variability in functional anatomy, this raises the question of how well the results can be generalized to the general population. The single case reported by Paul Broca in his pioneering study is a good example of this; in subsequent research, it has been observed that similar symptoms can be elicited after lesions that are highly variable within and around the areas that were originally implicated (Ojemann, 1979).

In animal studies, lesions to specific brain areas can be produced, and with modern techniques the lesions can be highly specific, targeting very small populations of neurons (Bajo et al., 2010). Naturally, however, ethical issues prevent induction of lesions in humans – with the exception of transient stimulation and disturbance of local brain function using transcranial magnetic stimulation (Barker et al., 1985), a method that will be described later in this chapter. A further challenge to studies of cognitive functions in brain-damaged patients is posed by brain plasticity. Behavioral / neuropsychological tests typically cannot be administered to brain damaged patients until several weeks after the has damage occurred. During this time, functional reorganization (*i.e.*, plasticity) has already taken place in the brain that easily biases the results that are obtained (for more on brain plasticity, see Chapter 8).

The double-dissociation method

The method of double dissociation has been an especially useful approach in neuropsychology. In this approach, two-way dissociation of functional deficits is sought to show that seemingly similar functions are carried out by different neural circuitry / mechanisms. For instance, if one manages to document a patient who, after brain damage, loses his/her ability to write, but can still read (constituting what is called a dissociation), and yet further manages to find another patient who has lost his/her ability to read but can still write, this constitutes a double dissociation of reading and writing abilities. A double dissociation of cognitive deficits strongly suggests that the functions are enabled by distinct neuronal mechanisms, and localization of the specific lesion sites provides information about the brain areas that are important for the functions in question.

Behavioral measures in healthy volunteers

Behavioral measures of performance, such as reaction times and number of errors committed, as well as various self-report questionnaires reveal important information not only about cognitive and perceptual processes, but also about underlying neural mechanisms. Of course, what cannot be done based on measuring performance alone is localization of function, however, that is only a part of what cognitive neuroscientists are interested in. As an example of this, consider that one wants to study the neural basis of empathy, the ability to feel for the other person. As will be described later in Chapter 11, there are distinct types of empathy: cognitive empathy and emotional empathy. Based on studies with neurological patients, it seems that different brain areas enable these two types of empathy (Shamay-Tsoory et al., 2009), however, without prior behavioral work where correlation patterns between questionnaire items were analyzed to dissociate these subtypes of empathy (Davis, 1983), one would not have known to look for and measure them separately in the patients.

With the behavioral measures a vast amount of information on the principles governing perceptual and cognitive functions have been obtained. For instance, most of what we know about the memory systems – that there are short-lived sensory memory traces, limited-capacity short-term memory, and long term memory (as will be described in more detail in Chapter 7), has been based on carefully designed behavioral experiments. One of the basic tests that have been used to infer the capacity of the short-term memory store consists of simply presenting the experimental subject series of digits that gradually increase in length, followed by immediate recall of what was presented. Initially there could be only 2 digits to remember, but as the test continues, the digit series increase in length so that there are as many as 10 digits in a string that need to be recalled. You can try this out yourself by writing down on pieces of paper random sequences of numbers of different lengths, picking them up one at a time and testing whether you are able to recite them after a single reading. Typically test subjects can recall 7 ± 2 digits, which is the “normal” capacity of the short-term memory.

As another example of a behavioral task that is used to measure cognitive functions, the Stroop color-word interference task (Stroop, 1935) is a widely used test of selective attention / ability to suppress learned responses (see **Figure 2-4**). In the Stroop test, the experimental subject is presented a list of color names that are written with different colors. The task of the subject is then to read aloud the color with which the color name is written. You can try this yourself, and probably notice the attentional effort that is needed to suppress the over-learned response to read aloud what is written. There are a large number of variants of this test that are used in cognitive neuroscience research today (MacLeod, 1991).

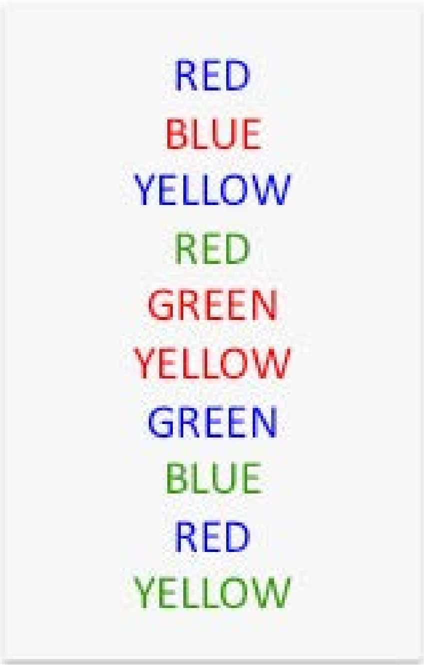


Figure 2-4. In the Stroop color-word interference task, the task of the subjects is read through the list aloud, so that the color with which a given color word is written is to be named. In order to accomplish this, the subject needs to suppress the automatic response to read aloud the color names. This task is used in many contexts to measure selective attention / ability to inhibit competing responses.

Behavioral tasks in animal models

In animal models, various behavioral tests have been designed, for instance, simple mazes that experimental animals need to learn. Experimental manipulations, such as administering certain pharmacological agents to the animals, can then be used to probe the role of specific neurotransmitter systems in spatial learning and navigation, by testing how the different pharmacological substances impact the maze learning ability. Even though there is no one-to-one correspondence between results obtained in animal and human studies, animal research is often invaluable in that certain tests cannot be conducted in humans due to ethical and/or legislative reasons (for instance, only specific medically approved drugs can be utilized to modulate neurotransmitter systems in humans), and findings from animal studies often guide research that is conducted in human volunteers by helping scientists to formulate specific testable hypotheses.

2.2 Invasive neuronal recordings in animal models

One of the most prominent methods that is used in animal models, and which cannot be used in humans except in certain patient cases during neurosurgery, is recording of neural activity with electrodes placed in the brain tissue. There are multiple ways in which neural activity can be recorded invasively in animal models, from recordings of single cells to recording of population activity. Animal models can be further divided into *in vivo* and *in vitro* studies, where the former means recording from living animals (which can be conducted under anesthesia or, increasingly, in awake animals that naturally makes the results more comparable with those obtained in human neuroimaging experiments) and the latter from tissue samples. *In vitro* tissue samples can be stimulated electrically to reveal functional characteristics and connectivity of different types of neurons within the sample. While studies conducted in live awake animals that are engaged in specific tasks and presented with stimuli are the most relevant for cognitive neuroscience, studies conducted *in vitro* or under anesthesia in living animals have contributed valuable information to our understanding of how the brain functions. The methods that are most commonly used *in vivo* are briefly described in the following.

Recording of single-unit activity

Neuronal firing (also called action potentials) can be recorded from single neurons *in vivo* and the relationship between neuronal firing and stimulus that is presented can be inspected. David Hubel and Torsten Wiesel published in the 1960s one of the classical observations using this procedure (see **Figure 2-5**). What Hubel and Wiesel discovered were visual cortex cells that respond to specific visual stimuli by increasing their firing when the respective stimulus is presented to the receptive field of the neuron (*i.e.*, the part of the visual field which' stimulation the neuron is sensitive to); the so-called simple cells reacted to lines of a specific orientation that were presented to the receptive field of the neuron (Hubel and Wiesel, 1959). Higher up in the visual hierarchy, cells started to react to more complex visual features, such as lines or edges of a specific orientation moving to a certain direction (Hubel and Wiesel, 1968). Hubel and Wiesel were awarded the Nobel Prize for their findings that still today form the basis of numerous computational models of human visual system function.

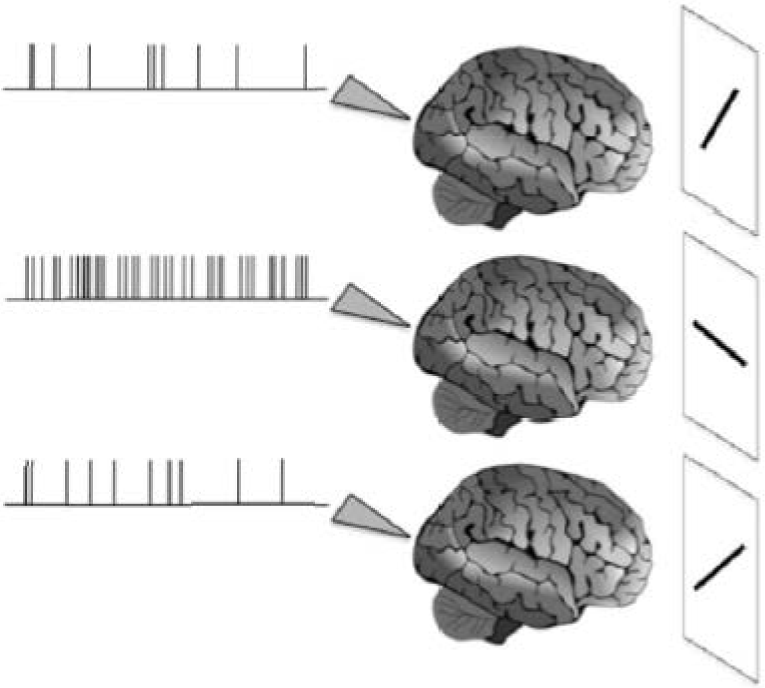


Figure 2-5. Schematic illustration of how a visual cortex simple cell preferentially responds to lines of a specific orientation presented in a part of the visual field that the cell is sensitive to (*i.e.*, the stimulus falls within the cell's receptive field). As can be seen, the cell increases significantly it's spiking rate (the vertical lines on the axes at left each represent a single spike or action potential) once a line of specific orientation falls within it's receptive field. The arrow points at the back of the head, to occipital cortex, wherein the primary visual cortex is buried in a fissure between the cerebral hemispheres. The simple cells were first documented in cats.

Multi-unit recordings and local field potentials

In addition to recording firing of single neurons (which can be tedious work, given that only one neuron can be recorded at a time and thus studies typically only report findings from a few tens of neurons) the collective activity of populations of neurons during various tasks and presentation of stimuli can be recorded in multi-unit recordings and by measuring local field potentials. These types of recordings are useful in that the average response of a given highly localized area can be associated with the tasks/stimuli. Furthermore, these population level responses seem to correlate with human neuroimaging measures more closely than single-unit recordings (Logothetis, 2002). The non-invasive human neuroimaging methods will be described in detail later in this chapter.

Optical imaging and two-photon calcium imaging techniques

In addition to recording the electrical activity of neurons using electrodes, there are sophisticated methods whereby neuronal activity can be assessed even *in vivo* optically with electrosensitive dyes (Grinvald et al., 1984). These dyes change color voltage-dependently, which can be quantified by the means of optical measurements. Optical imaging thus offers a way to inspect electrical activity of populations of neurons in animal models. Closely related to the optical imaging methods, the so-called two-photon calcium imaging techniques (Ohki et al., 2005) are emerging as a highly prominent method where fluctuations in the level of intracellular calcium can be simultaneously assessed from populations of thousands or even tens of thousands individual neurons (see **Figure 2-6**). This is based on injecting specific marker substance to the imaged area that makes it possible to detect intracellular calcium levels. As will be described in detail in Chapter 4, intracellular calcium of a neuron increases when the neuron becomes more active in sending neural impulses to other neurons.

The main difference between optical imaging and two-photon imaging methods is that with two-photon imaging, it is possible to measure simultaneously the activity of each neuron in a large population of neurons, whereas optical imaging (as well as the multi-unit and local field potential measures) only allows inspection of the responses of the neuronal population as a whole. Thus, the two-photon imaging method effectively allows inspection of network activity of individual neurons. Further, the functional images obtained with two-photon calcium imaging can be overlaid with highly accurate structural images obtained with electron microscopy to study structure-function relationships at the level of single neurons that make up the larger populations (Bock et al., 2011).

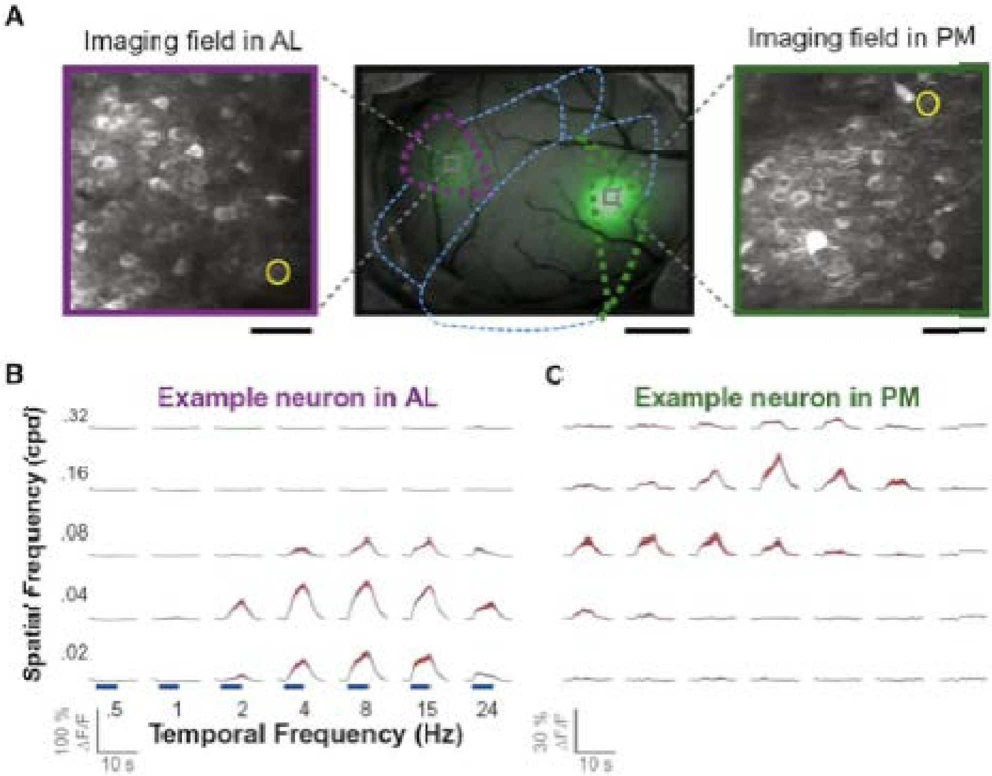


Figure 2-6. Shown is an example of two-photon calcium imaging results that disclose functional specialization of single neurons in distinct mouse visual cortical areas. A) Mouse anterolateral (AL) and posteromedial (PM) visual areas are shown enlarged (scale bars 50 µm on the left and right panels, and 1 mm in the middle panel that shows the mouse visual areas) with exemplary individual neurons encircled. B) Averaged responses as a function of spatial and temporal frequency of visual stimulation from the example neuron in area AL, C) Averaged responses as a function of spatial and temporal frequency of visual stimulation from the example neuron in area PM. As can be seen comparing B vs. C, these neurons respond in differential manner to spatial and temporal aspects of visual stimulation. Adapted from (Andermann et al., 2011).

2.3 Non-invasive functional neuroimaging methods

While animal models are highly useful, perceptual and cognitive functions are to a large extent specific to the human species. For instance, language functions exits almost exclusively in humans. Thus, the ability to record human brain activity is of utmost significance for cognitive neuroscience. Fortunately, there are several non-invasive functional neuroimaging methods that can be utilized to measure various aspects of brain activity in humans. In fact, the rapid progression in the methods that are available, improvements in the signal-to-noise ratio of these methods, as well as advances in data analysis methods that make sense of the complex signals obtained, play a pivotal role in cognitive neuroscience. In the following the most common non-invasive neuroimaging methods are briefly introduced.

Electroencephalography and event-related potentials

Electroencephalography (EEG) is a relatively old method with the first EEG recordings documented already in 1924 by German scientist Hans Berger. EEG measures electric potential differences, caused by currents generated by synchronous activity (*i.e.*, post-synaptic potentials) of large populations of so-called cortical pyramidal neurons, with electrodes attached on the scalp of the subjects. Whereas in early studies only few electrodes were typically used, modern EEG systems come with caps with 64 or even up to 256 electrodes. When recording EEG, one of the electrodes serves as a reference electrode against which the potential difference is measured as a function of time.

High temporal resolution (*i.e.*, milliseconds) is the clear advantage of the EEG. Early EEG work documented so-called brain rhythms that are still in use in clinical settings today: in EEG spectra there are peaks in delta (up to 4 Hz) theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–100 Hz) frequencies (the frequency in Hz here means fluctuation cycles per second and a Fourier transform is used to obtain the spectra from EEG time series). During the last couple of decades, interest towards these rhythms has increased also in cognitive neuroscience, as associations between rhythmic/oscillatory activity and network properties of brain activity have been documented that appear to be relevant for perceptual and cognitive functions.

The gamma-band activity has especially generated interest, as it has been associated with the so-called "binding problem", specifically, how brain areas processing different aspects of the same stimulus (*e.g.*, the red color, texture and shape of an apple) are bound together to give rise to a coherent percept (Singer and Gray, 1995, Bertrand and Tallon-Baudry, 2000). It has been hypothesized that different brain areas synchronize to the same phase of gamma frequency and that this phase synchrony helps bind parallel processing that takes place across the distinct brain areas into a perceptual object.

The high temporal resolution of the EEG also allows one to inspect the responses of the brain to various stimuli under different task conditions. Since the stimulus-related responses are much smaller than the spontaneous rhythms of the brain, this is typically made possible by repeating the stimulus tens or even hundreds of times, assigning a trigger to each stimulus event and then simply averaging the EEG epochs time-locked to the onset of the stimuli. The averaging process helps reduce noise and preserve the signal, thus revealing what is called an event-related potential (ERP) (see **Figure 2-7**).

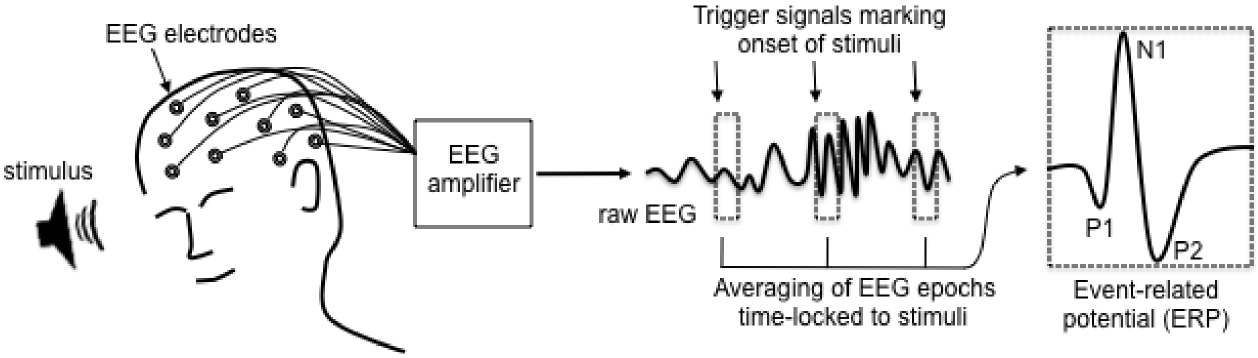


Figure 2-7: Electroencephalography (EEG) based recording of event-related potentials (ERP). In this schematic illustration, from left to right, sound stimuli are presented to a subject and EEG electrodes attached to the scalp of the subject record electric potential differences as a function of time with millisecond accuracy. The signals are amplified and from the resulting "raw EEG" epochs time-locked to onset of stimuli (as marked by specific triggers) are averaged, yielding the event-related potential that is composed of several deflections that are named based on whether they are negative or positive (N/P) and their order. Here are plotted the most prominent deflections, P1, N1, and P2.

There are several deflections in the ERP that are denoted based on whether they are negative or positive in polarity (note that by convention, negative is plotted up and positive down) and by their latency or serial order. For example, perhaps the most prominent auditory event-related potential deflection that is negative in polarity and occurs at about 100 ms from sound onset is called either N100 or N1, since it is the first robust negative-polarity deflection, even though it does not reflect a unitary cerebral event but is contributed to by multiple processes (Näätänen and Picton, 1987). While there are negative-going deflections that do occur prior to the N1 called middle-latency responses, these are much smaller in magnitude and in the early days of EEG they were difficult to discern even from the averaged EEG. Thus, the response at about 100 ms from stimulus onset has been named the N1.

With stimulus and task manipulations, it is possible to modulate the event-related potential waveform. For example, when the task of the subjects is to detect specific target stimuli that are embedded amongst non-target stimuli, there is a late positive-polarity response at about 300 ms from target sound onset called P300 (Donchin and Cohen, 1967). Such task-dependent responses are also called endogenous responses, to dissociate them from the exogenous responses that presumably occur invariantly to every stimulus regardless of the task. However, this should not be seen as a categorical difference, as for example the N1 that has been viewed as an exogenous response that is modulated by task requirements, for example, selective attention augments the N1 response (Hillyard et al., 1973).

Substantial effort has been devoted to development of methods that can be used to estimate where the electric potentials measured at the scalp are generated in the brain. It is good to realize that there is always ambiguity in inverse estimation, given that the scalp-recorded potentials can be explained by infinite number of different underlying source configurations (Hämäläinen et al., 1993). However, when EEG (and also magnetoencephalography; MEG, see below) inverse estimates have been compared to what is known about functional anatomy of the brain based on other studies, it can be concluded that the inverse estimates do seem to be fairly accurate. Further, using *a priori* information such as knowledge of the anatomy based on MRIs and loci of hemodynamic activity, the accuracy of the electromagnetic inverse estimates can be further increased (Liu et al., 1998, Auranen et al., 2009).

Magnetoencephalography and event-related magnetic fields

MEG is a method that was developed several decades after the invention of EEG (Hämäläinen et al., 1993). Similarly to EEG, MEG has a very good temporal resolution of milliseconds. MEG measures extremely weak magnetic field changes that are generated by electric currents within the brain. The magnetic field changes induce weak currents to superconductive MEG sensor elements that are placed close to the surface of the head. These currents that are then quantified as a function of time. As is also the case in recording electric potentials with EEG, synchronous activity (*i.e.*, post-synaptic potentials) of thousands of pyramidal neurons is required in order to generate currents that are strong enough to be picked up by the MEG sensors.

Given the geometry of the head that resembles a sphere, MEG detects mainly magnetic fields generated by the tangential components of the source currents. Further, given that the cortical pyramidal cells are oriented perpendicular to the cortical surface and that the magnetic field decays rapidly as the distance between the sensor and the source increases, MEG is more sensitive to source currents that are close to the surface of the brain and located within the sulci of the cortical foldings (Hillebrand and Barnes, 2002). Sources located deeper in the brain are more difficult to measure using MEG than EEG, given that the magnetic field rapidly weakens as distance to the sensors grows.

One of the advantages provided by MEG is that the current sources are easier to localize in the brain using MEG than EEG. This is due to the skull not smearing the magnetic fields (unlike the electric potentials that are measured using EEG). It has been, however, shown that the best source localization results are obtained when information provided by simultaneous recording of MEG and EEG is combined using, for instance, so-called linear inverse estimates (Liu et al., 2002).

Positron emission tomography

Positron emission tomography (PET) is based on rather different principles than EEG or MEG. Prior to invention of functional magnetic resonance imaging (fMRI; discussed below), cognitive activation studies were done mostly using PET (Cabeza and Nyberg, 1997). In PET, weakly radioactive substance is injected to the circulation of the subject. As this tracer substance circulates through the brain, it releases positrons that, when colliding with electrons, become energy that can be detected outside of the head with so-called co-incidence detectors. With arrays of such detectors, a three-dimensional picture can be reconstructed of where in the brain the tracer substance is concentrated.

Depending on the type of tracer substance, PET can be used to study distribution of different types of receptors (*i.e.*, when a tracer is used that binds to specific receptors; for description of receptor see Chapter 4) in the brain, and also blood flow can be quantified when labeled oxygen is used as a tracer. Since blood flow and oxygen delivery increases to brain areas that are more active, this allows one to detect brain areas that participate in processing specific stimuli and in different perceptual and cognitive tasks that the subjects are engaged in during imaging.

While the PET does not suffer from the ill-posed inverse problem of EEG and MEG, and localization of the radiation sources is highly accurate, PET has a very limited temporal accuracy (from tens of seconds to minutes) making it very difficult to inspect dynamic activity of the brain. Also, since the participants are subjected to weak radiation during PET studies, a given subject cannot, for strict safety reasons, participate more than once in a PET study. This, along with relatively short half-lives of some of the tracer substances sets some limits to the scope of PET studies.

Functional magnetic resonance imaging

Functional magnetic resonance imaging (fMRI) is a relatively new, and already perhaps the most widely used, method that allows spatially accurate (millimeters) quantification of blood oxygen level changes in the brain (Ogawa et al., 1990, Belliveau et al., 1991). This is based on differential magnetic properties of red blood cells (hemoglobin) when they carry oxygen (oxyhemoglobin) *vs.* when depleted of oxygen (deoxyhemoglobin). While oxyhemoglobin is diamagnetic, deoxyhemoglobin is paramagnetic. In order to understand how this difference shows up in fMRI, one needs to understand some of the basics of magnetic resonance. In fMRI, the head of the subject is placed in a very strong homogeneous magnetic field, with the field strength varying, depending on the scanner, between 1.5 – 7 Tesla. This strong magnetic field aligns a portion of atomic nuclei to rotate along the magnetic field. The frequency of this rotation, called the Larmor frequency, depends on the magnetic field strength. A radiofrequency pulse applied at the Larmor frequency pushes the rotating spins off from their alignment with the main magnetic field. As the radiofrequency pulse is turned off, the rotating spins start to realign with the magnetic field, at the same time releasing energy that can be picked up by so-called readout coils.

By applying gradients to the homogeneous magnetic field, one can select specific portions of the imaged object to be sensitive

to the radiofrequency pulse, and one can also alter the rotation speed (and thus the frequency of the readout signal) of spins in selected parts of the imaged object. This way, the energy released by the spins after the radiofrequency pulse can be located to specific parts of the imaged object and a three-dimensional picture of the brain can be reconstructed. With specifically designed MRI sequences where gradients switch very rapidly it is possible to image the whole brain in a couple of seconds. The ratio between oxyhemoglobin and deoxyhemoglobin affects the local signal strength in the resulting fMRI images, as the paramagnetic deoxyhemoglobin creates local magnetic field inhomogeneity, thus reducing signal. Then, as neuronal populations increase their activity, diamagnetic oxygenated blood rushes into the area and increases the local magnetic field homogeneity, thus resulting in MR signal increase.

The MR signal intensity changes that are caused by activity-dependent variations in blood oxygenation level are rather small, on the order of 1-4% in 1.5 Tesla systems, and therefore statistical data analysis methods are needed to determine which signal changes occur consistently throughout the experiment, and which can be discarded as noise. In the simplest case, visual stimulation could be contrasted with a baseline condition where no visual stimuli are presented. By alternating between visual stimulation and no-stimulation baseline conditions, it is possible to calculate for each element of the three-dimensional volume (*i.e.*, voxel) distributions of signal intensity during stimulation *vs*. baseline (Kwong et al., 1992). Then, based on the degree of overlap between the two distributions, it can be statistically estimated whether two distributions differ significantly from each other. For a figure illustrating this type of approach in one of the first functional mapping experiments using fMRI, see **Figure 2-8**.

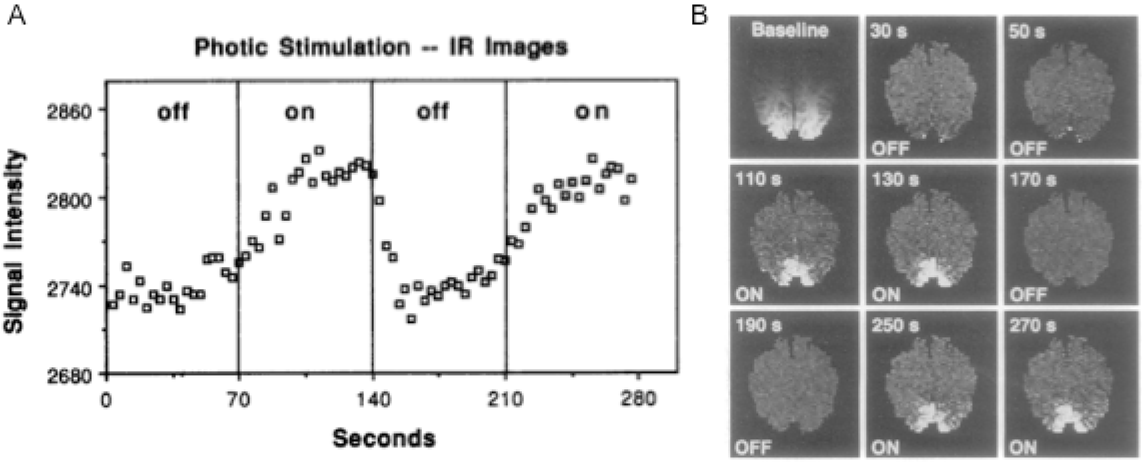


Figure 2-8. Visual cortex activation during photic stimulation *vs*. darkness in one of the first fMRI experiments. A) MR signal intensity in the visual cortex is plotted as a function time across the conditions where photic stimulation has been turned off and on. As can be seen, the MR signal intensity increases in the visual cortex when the photic stimulation is turned on, and decreases in darkness. B) The same effect seen in the fMRI images. Here, images are subtracted from the baseline image shown at the upper left hand corner. The MR signal intensity changes caused by the photic stimulation can be clearly seen in the subtraction images. Adapted from (Kwong et al., 1992).

In modern fMRI experiments, more sophisticated data analysis methods are typically used and software packages are available in the worldwide web, such as FSL (www.fmrib.ox.ac.uk/fsl/) and SPM (www.fil.ion.ucl.ac.uk/spm/). Recently, advances in signal analysis methods have even enabled use of highly naturalistic stimuli such as feature movies (Hasson et al., 2004, Jaaskelainen et al., 2008) and computer games (Spiers and Maguire, 2006) as stimuli in fMRI studies. For an exciting example of the use of a computer game in neuroimaging studies, see **Figure 2-9**. Furthermore, the spatially accurate information provided by fMRI has also been used as *a priori* information when solving the MEG (and EEG) inverse problem. This type of combination of methodology has been shown to significantly increase the accuracy of source localization (Liu et al., 1998).

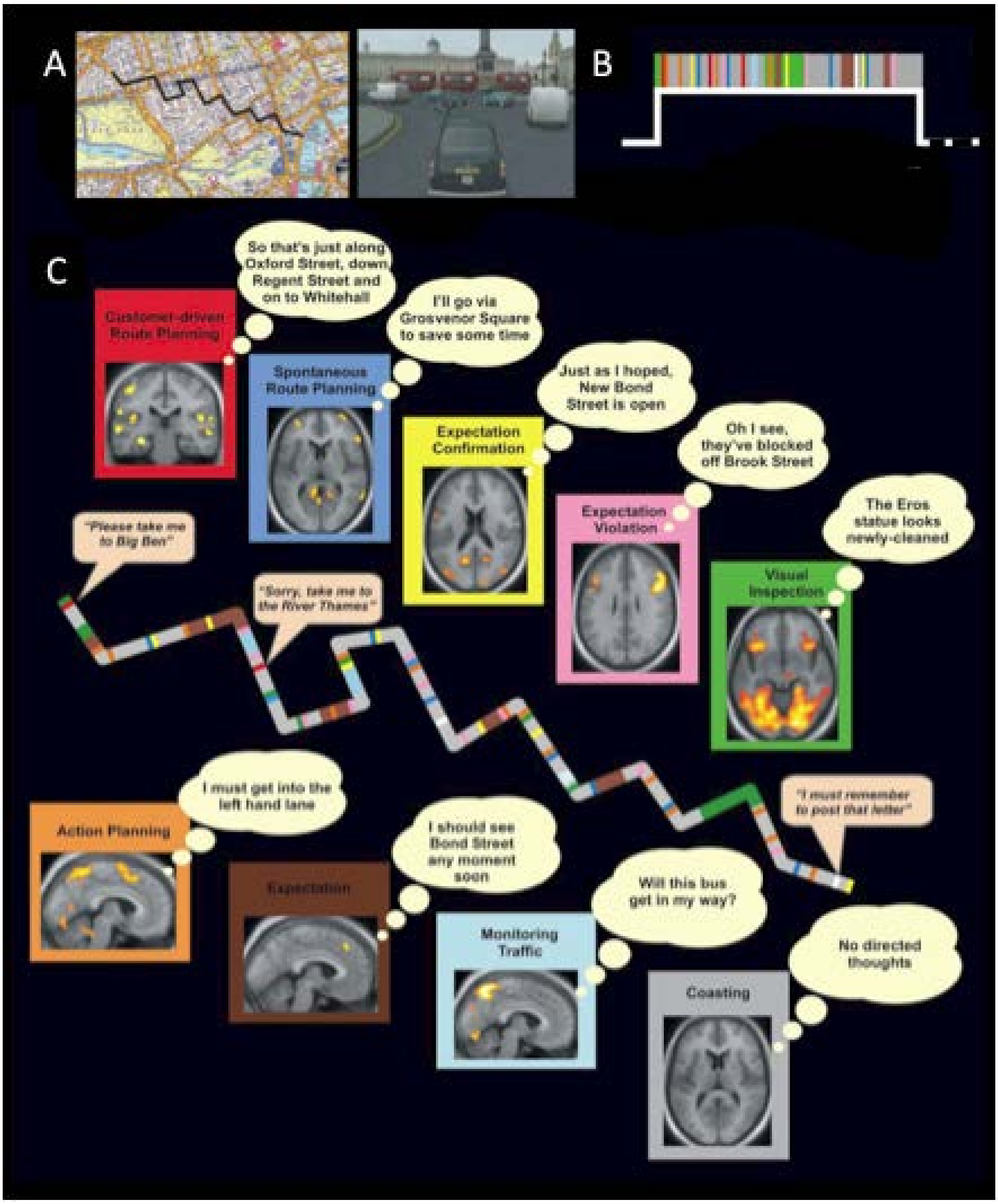


Figure 2-9. Shown is an example of a modern fMRI study where naturalistic stimulus/task was used. A) Here, the brain activity of participants (in this case London cab drivers) was recorded while they played a computer game in which the task was to navigate a cab on the streets of London. B) A time series of various thought patterns during specific game events, obtained afterwards with self-report questionnaires, was constructed to enable statistical analysis of the data. C) These events were associated with patterns of brain activity to reveal sets of brain areas activated during, for instance, traffic monitoring and route planning. Adapted from (Spiers and Maguire, 2006).

Dynamic inverse imaging: fMRI with faster temporal sampling rates

The relatively slow sampling rate that is achieved using conventional fMRI sequences (depending on the scanner and sequence used, one whole-brain volume is typically obtained every 1-3 seconds) has been a limitation for fMRI. A recently developed fMRI technique called dynamic inverse imaging (InI) allows an order of magnitude faster sampling rate (Lin et al., 2006). In conventional fMRI, the speed of the scanning procedure is limited by the need to switch magnetic field gradients in order to decode which parts of the imaged object the signals picked up by the coils emanate from. The rapid gradient switching also produces acoustic noise, in some scanner models up to even 140 dB, that easily becomes a confound in cognitive and perceptual studies. In contrast, InI utilizes an array of receiver coils to decode the origin of the MR signal in the imaged object, based on the differential spatial sensitivity profiles of the different receiver coils. Slightly reduced spatial accuracy is the downside of InI, and it is difficult to implement in older fMRI scanners that do not have capability to use large enough receiver coil arrays. However, the InI clearly holds promise for being one of the most significant neuroimaging methods in the near future.

Near-infrared spectroscopy

Near-infrared spectroscopy (NIRS), also often referred to as optical imaging (even though one has to be careful not to confuse it with the invasive optical imaging methods that are used in animal models described above), is a relatively new and rapidly developing method that, similar to fMRI, measures changes in the oxygen concentration of blood (Kato et al., 1993). Rather than differential magnetic properties, optical imaging is based on color differences between oxygenated and deoxygenated blood (see **Figure 2-10**). In NIRS, near-infrared light is transmitted to the imaged object and specific sensors pick up light that scatters back through the object. You can try this by shining light on your finger with a flashlight or a laser pointer in a dark room: you can see how some of the light is emitted from the side of your finger. Should the oxygenation level of blood circulating in your finger change, the spectra of the light scattered back would change and this could be measured with specific sensors.

The advantages offered by NIRS are the relatively small size of the imaging equipment (compared with the MRI scanners that weight thousands of kilograms and require massive magnetically shielded rooms) and lack of acoustic noise (that the rapid gradient switching produces during conventional fMRI). One of the shortcomings of NIRS is that the method does not detect sources of activity deeper in the sulci of the cortex, or in midbrain structures, however, NIRS has established itself as one of the more prominent non-invasive neuroimaging methods and seems to be especially useful in imaging brain activity of infants, where the smaller size of the head and larger permeability of the skull to near-infrared electromagnetic radiation enhances the applicability of the method. Furthermore, the complete non-invasiveness of the NIRS method makes it ethically sound for infant studies where use of, for instance, PET with the radioactive tracers would not be possible.

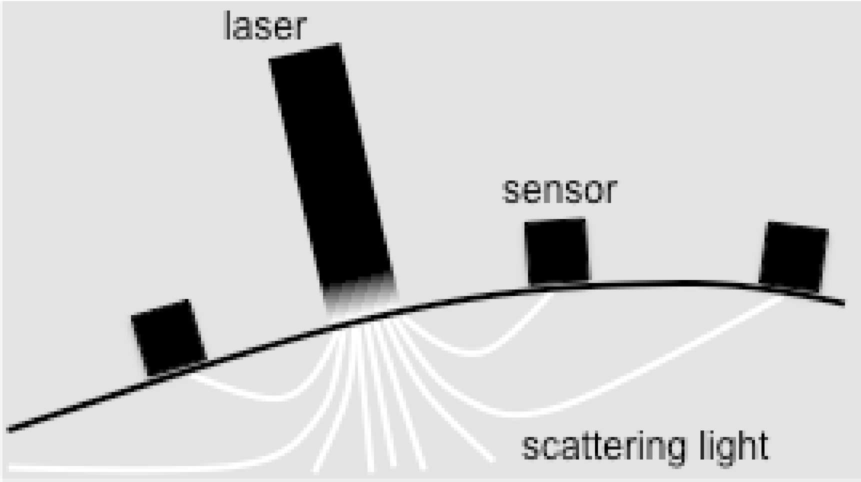
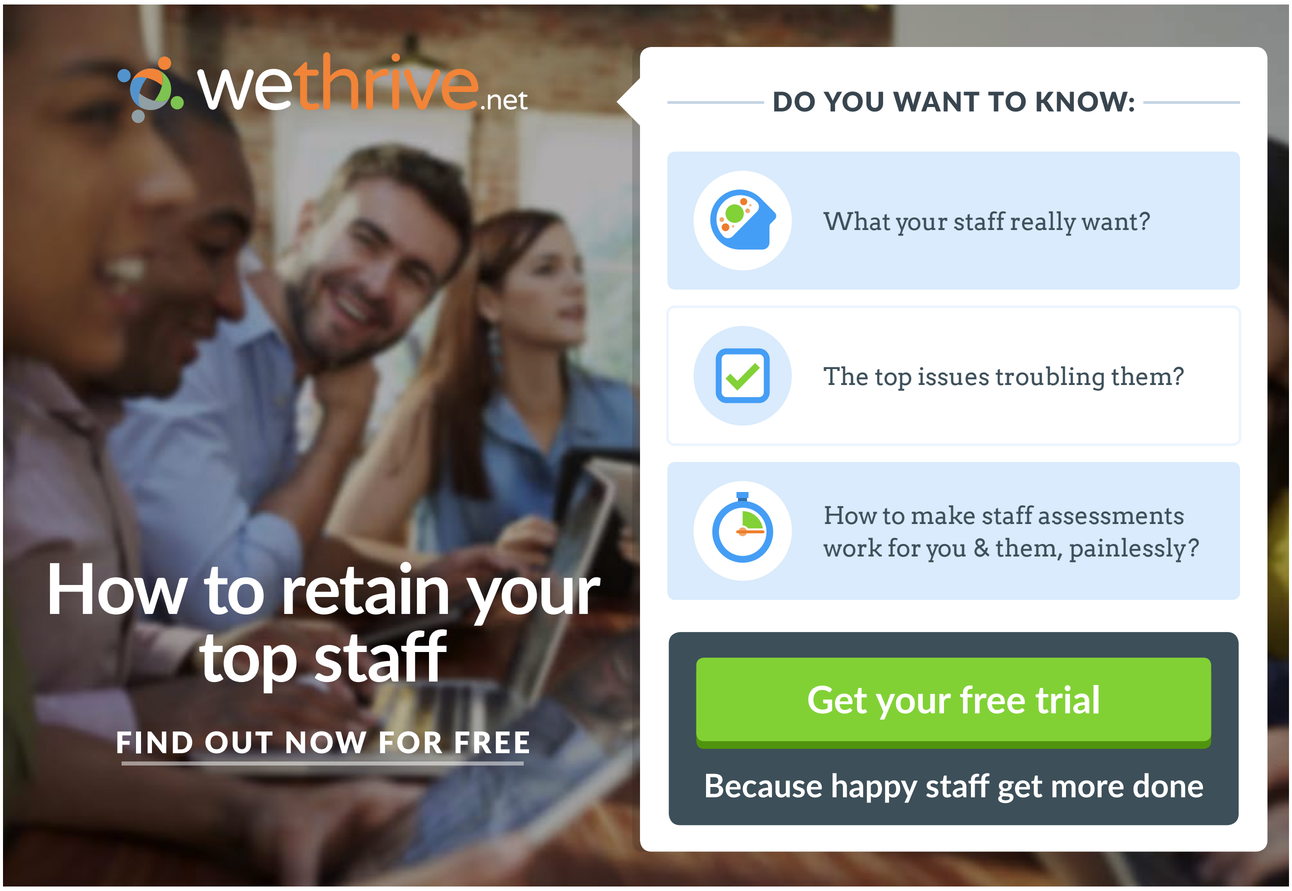


Figure 2-10. The operating principle of NIRS (aka optical imaging) is shown in this schematic illustration. Electromagnetic radiation (*i.e.*, light) is produced at near infrared frequencies with a laser and shined through the imaged object. Some of the light scatters and returns back to surface near the light source, where the sensors are located. The spectra of light change as the optical properties of the imaged object that the scattered light travels through change. Based on this, the sensors can detect blood oxygen concentration changes, given the differential light-absorbing properties (*i.e.*, color) of oxygenated and deoxygenated hemoglobin.



Transcranial magnetic stimulation

Transcranial magnetic stimulation (TMS) is a method that is in some ways the reverse of MEG. Whereas MEG picks up the the extremely weak magnetic field changes produced by neuronal currents, in TMS strong currents are fed into coils placed on the surface of the head that give rise to a powerful magnetic field. The magnetic field then induces currents in the brain that can, depending on the stimulation paradigm, either activate or inhibit neuronal populations in localized cortical areas on the surface of the brain (Barker et al., 1985, Pascual-Leone et al., 2000).

Using TMS, it is possible to test whether a given cortical area plays a role in a specified perceptual/cognitive function by stimulating the brain during task performance. For instance, stimulating the hand area of the primary motor cortex results in involuntary twitching of the hand as the currents induced in the motor cortex result in contraction of the hand muscles. Often this is used in TMS studies as a measure to set the stimulation threshold so that the applied pulse is just strong enough to induce movements (*i.e.*, “motor threshold”). Use of TMS targeted to the Broca’s area in testing the hypothesis that Broca’s area is specifically involved in syntactic processing of language is an example of a study where TMS has been used to clarify the neural basis of higher cognitive functions (Sakai et al., 2002).

TMS has also been successfully combined with other non-invasive neuroimaging methodology. Anatomical information provided by MRI has been used to guide the targeting of TMS pulses to specific structures that have been hypothesized as functionally relevant based on prior knowledge of functional anatomy or separately acquired fMRI scans (Gugino et al., 2001). Additionally, recording of EEG has been coupled with TMS to study how the TMS pulses affect electrical activity of the brain (Izumi et al., 1997).

2.4 Concluding remarks

Development of research methodology, especially non-invasive structural and functional neuroimaging methods, plays a pivotal role in cognitive neuroscience. Independently, each of the methods is compromised in spatial or temporal resolution. When information across methods is combined, however, relatively high spatiotemporal resolution can be presently achieved. The modern neuroimaging techniques and associated sophisticated data analysis methods make it possible for cognitive neuroscientists to provide answers to specific hypotheses that are formulated based on theoretical models, either proving or disproving (and thus refining) the theoretical models. Indeed, the possibilities offered by modern neuroimaging technology are quite amazing; still some decades ago, neuroscientists could have only dreamt about methods that allow one to image the function of the living brain non-invasively and associate cognitive processes with specific anatomical structures. To continue the introduction to cognitive neuroscience, the major anatomical structures of the brain will be introduced in the next chapter of this book.

2.5 References

ANDERMANN, M. L., KERLIN, A. M., ROUMIS, D. K., GLICKFELD, L. L. & REID, R. C. 2011. Functional specialization of mouse higher visual cortical areas. *Neuron*, 72, 1025-1039.

AURANEN, T., NUMMENMAA, A., VANNI, S., VEHTARI, A., HÄMÄLÄINEN, M., LAMPINEN, J. & JÄÄSKELÄINEN, I. P. 2009. Automatic fMRI-guided MEG multidipole localization for visual responses. *Hum Brain Mapp,* 30**,** 1087-1099.

BAJO, V. M., NODAL, F. R., MOORE, D. R. & KING, A. 2010. The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nature Neuroscience,* 13**,** 253-260.

BARKER, A. T., JALINOS, R. & FREESTON, I. L. 1985. Non-invasive magnetic stimulation of human motor cortex. *Lancet,* 1**,** 1106–1107.

BELLIVEAU, J. W., KENNEDY, D. N., MCKINSTRY, R. C., BUCHBINDER, B. R., WEISSKOFF, R. M., COHEN, M. S., VEVEA, J. M., BRADY, T. J. & ROSEN, B. R. 1991. Functional mapping of the human visual cortex by magnetic resonance imaging. *Science,* 254**,** 716-719.

BERTRAND, O. & TALLON-BAUDRY, C. 2000. Oscillatory gamma activity in humans: a possible role for object representation. *International Journal of Psychophysiology,* 38**,** 211-223.

BOCK, D. D., LEE, W. C. A., KERLIN, A. M., ANDERMANN, M. L., HOOD, G., WETZEL, A. W., YURGENSON, S., SOUCY, E., KIM, H. S. & REID, R. C. 2011. Network anatomy and in vivo physiology of visual cortical neurons. *Nature,* 471**,** 177-182.

BROCA, P. 1861. Remarque sur le siege de la faculté du language articulé, suivie d'une observation d'aphémie (perte de la parole). *Bulletin de la société anatomique de Paris,* 36**,** 330-356.

CABEZA, R. & NYBERG, L. 1997. Imaging cognition: an empirical review of PET studies with normal subjects. *Journal of Cognitive Neuroscience,* 9**,** 1-26.

DAVIS, M. H. 1983. Measuring individual differences in empathy: evidence for a multidimensional approach. *J Pers Soc Psychol,* 44**,** 113-126.

DONCHIN, E. & COHEN, L. 1967. Averaged evoked potentials and intramodality selective attention. *Electroencephalography and Clinical Neurophysiology,* 22**,** 537-546.

GRINVALD, A., ANGLISTER, L., FREEMAN, J. A., HILDESHEIM, R. & MANKER, A. 1984. Real-time optical imaging of naturally evoked electrical activity in intact frog brain. *Nature,* 308**,** 848-850.

GUGINO, L. D., ROMERO, J. R., AGLIO, L., TITONE, D., RAMIREZ, M., PASCUAL-LEONE, A., GRIMSON, E., WEISENFELD, N., KIKINIS, R. & SHENTON, M. E. 2001. Transcranial magnetic stimulation coregistered with MRI: a comparison of a guided versus blind stimulation technique and its effect on evoked compound muscle action potentials. *Clin Neurophysiol,* 112**,** 1781-1792.

HÄMÄLÄINEN, M., HARI, R., ILMONIEMI, R. J., KNUUTILA, J. & LOUNASMAA, O. V. 1993. Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev mod phys,* 65**,** 413-497.

HASSON, U., NIR, Y., LEVY, I., FUHRMANN, G. & MALACH, R. 2004. Intersubject synchronization of cortical activity during natural vision. *Science,* 303**,** 1634-1640.

HILLEBRAND, A. & BARNES, G. R. 2002. A quantitative assessment of the sensitivity of whole-head MEG to activity in the adult human cortex. *Neuroimage,* 16, 638-650.

HILLYARD, S. A., HINK, R. F., SCHWENT, V. L. & PICTON, T. W. 1973. Electrical signs of selective attention in the human brain. *Science,* 182, 177-180.

HUBEL, D. H. & WIESEL, T. N. 1959. Receptive fields of single neurons in the cat's striate cortex. *Journal of Physiology,* 148, 574-591.

HUBEL, D. H. & WIESEL, T. N. 1968. Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology,* 195, 215-243.

IZUMI, S., TAKASE, M., ARITA, M., MASAKADO, Y., KIMURA, A. & CHINO, N. 1997. Transcranial magnetic stimulation-induced changes in EEG and responses recorded from the scalp of healthy humans. *Electroencephalography and Clinical Neurophysiology,* 103, 319-322.

JAASKELAINEN, I. P., KOSKENTALO, K., BALK, M. H., AUTTI, T., KAURAMÄKI, J., POMREN, C. & SAMS, M. 2008. Inter-subject synchronization of prefrontal cortex hemodynamic activity during natural viewing. *The Open Neuroimaging Journal,* 2, 14-19.

KATO, T., KAMEI, A., TAKASHIMA, S. & OZAKI, T. 1993. Human visual cortical function during photic stimulation monitoring by means of near-infrared spectroscopy. *J Cereb Blood Flow Metab,* 13, 516-520.

KWONG, K. K., BELLIVEAU, J. W., CHESLER, D. A., GOLDBERG, I. E., WEISSKOFF, R. M., PONCELET, B. P., KENNEDY, D. N., HOPPEL, B. E., COHEN, M. S., TURNER, R., CHENG, H. M., BRADY, T. J. & ROSEN, B. R. 1992. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc Natl Acad Sci USA,* 89**,** 5675-5679.

LEZAK, M. D., HOWIESON, D. & LORING, D. W. 2004. *Neuropsychological assessment (4th ed.),* Oxford, Oxford University Press.

LIN, F.-H., WALD, L. L., AHLFORS, S. P., HAMALAINEN, M., KWONG, K. K. & BELLIVEAU, J. W. 2006. Dynamic magnetic resonance inverse imaging of human brain function. *Magn Reson Med,* 56**,** 787-802.

LIU, A. K., BELLIVEAU, J. W. & DALE, A. M. 1998. Spatiotemporal imaging of the human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations. *Proc Natl Acad Sci USA,* 95**,** 8945-8950.

LIU, A. K., DALE, A. M. & BELLIVEAU, J. W. 2002. Monte Carlo simulation studies of EEG and MEG localization accuracy. *Hum Brain Mapp,* 16**,** 47-62.

LOGOTHETIS, N. K. 2002. The neural basis of blood-oxygen-level-dependent functional magnetic resonance imaging signal. *Philos Trans R Soc Lond B Biol Sci,* 357**,** 1003-1037.

MACLEOD, C. M. 1991. Half a century of research on the stroop effect: an integrative review. *Psychological Bulletin,* 109**,** 163-203.

NÄÄTÄNEN, R. & PICTON, T. 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of component structure. *Psychophysiology,* 24**,** 375-425.

NEWCOMBE, F. & RUSSELL, W. R. 1969. Dissociated visual perceptual and spatial deficits in focal lesions of the right hemisphere. *J. Neurol. Neurosurg. Psychiat.,* 32**,** 73-81.

OGAWA, S., LEE, T. M., RAY, A. R. & TANK, D. W. 1990. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci USA,* 87**,** 9868-9872.

OHKI, K., CHUNG, S., CH'NG, Y. H., KARA, P. & REID, R. C. 2005. Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature,* 433**,** 597-603.

OJEMANN, G. A. 1979. Individual variability in cortical localization of language. *J Neurosurg,* 50**,** 164-169.

PASCUAL-LEONE, A., WALSH, V. & ROTHWELL, J. 2000. Transcranial magnetic stimulation in cognitive neuroscience--virtual lesion, chronometry, and functional connectivity. *Curr Opin Neurobiol,* 10**,** 232-237.

RHODES, G., JEFFERY, L., WATSON, T.L., JAQUET, E., WINKLER, C., CLIFFORD, C.W.G. 2004. Orientation-contingent face aftereffects and implications for face-coding mechanisms. *Current Biology*, 14, 2119-2123.

SAKAI, K. L., NOGUCHI, Y., TAKEUCHI, T. & WATANABE, E. 2002. Selective priming of syntactic processing by event-related transcranial magnetic stimulation of Broca's area. *Neuron,* 35**,** 1177-1182.

SEKULER, R. W., RUBIN, E. L. & CUSHMAN, W. H. 1968. Selectivities of human visual mechanisms for direction of movement and contour orientation. *J. Opt. Soc. Am.,* 58**,** 1146-1146.

SHAMAY-TSOORY, S. G., AHARON-PERETZ, J. & PERRY, D. 2009. Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain,* 132**,** 617-627.

SINGER, W. & GRAY, C. M. 1995. Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience,* 18**,** 555-586.

SPIERS, H. J. & MAGUIRE, E. A. 2006. Thoughts, behavior, and brain dynamics during navigation in the real world. *Neuroimage,* 31**,** 1826-1840.

STROOP, J. R. 1935. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology,* 18**,** 643-662.

WEBSTER, M. A., KAPING, D., MIZOKAMI, Y. & DUHAMEL, P. 2004. Adaptation to natural facial categories. *Nature,* 428**,** 557-561.

WERNICKE, C. 1874. *Der aphasische symptomencomplex: eine psychologische studie auf anatomischer basis,* Breslau, Cohn & Weigert.

3 Anatomy of the brain

In order to be able to study the neural basis of perceptual and cognitive functions, it is necessary to have some knowledge of brain anatomy (*i.e.*, what the names of different brain structures/areas are and where they are located with respect to one another), as cognitive neuroscience literature is to a large extent grounded on associating specific types of cognitive and perceptual functions and their constituent processes with various anatomical structures. Learning brain anatomy is made more complicated than what it could otherwise be due to certain brain areas having multiple names. For instance, the brain area that processes visual information can be called the primary visual cortex, striate cortex, and V1. Furthermore, in some cases there is disagreement on where the boundaries of different areas are. For instance, the most anterior tip of the brain called frontal pole is by some definitions just the very tip, and according to other definitions a relatively large part of, the frontal part of the brain. The reasons for this are partly historical with scientists using different samples and techniques arriving at slightly different definitions.

Indeed, what needs to be borne in mind is that two brains are structurally never identical. Even identical twins exhibit differences in brain structure, especially in gyral patterns (Bartley et al., 1997). Localization of function to anatomy is even more variable, than anatomical variability, across individuals. Furthermore, one needs to be cautious when anatomical structures are compared across species, as there are substantial functional and structural differences even though many of the structures carry identical names. In order to learn the basic terminology of brain anatomy, it is first necessary to learn the coordinate system(s) that are used in defining the loci of various brain structures.

3.1 Anatomical coordinate systems

Brain anatomical coordinate systems are the first potential point of confusion, as there are different ones (shown in **Figure 3-1** below) due to the fact that research is conducted in different species where different conventions have been used. One of the coordinate systems is based on naming areas based on whether they are located closer to the head of the animal (rostral) *vs.* the tail of the animal (caudal) and, as the other axis, whether the area is more towards the back (dorsal) or the abdomen (ventral). In humans, there is no straight line from the forehead to the "tail" but rather the human brain has evolved in a curved structure. As a result of this, the coordinate system is also curved as shown in **Figure 3-1**.

Another coordinate system that is used in cognitive neuroscience has anterior-posterior as one and superior-inferior as the other axis. This coordinate system is not curved. In both coordinate systems, there is a third coordinate that runs from one ear to the other, and brain structures are further defined based on whether they are on the left or the right side of the midpoint between the ears. To define relative positions of brain structures along this dimension, the terms lateral (closer to the surface/ear) and medial (deeper in the brain/closer to midline) are used.

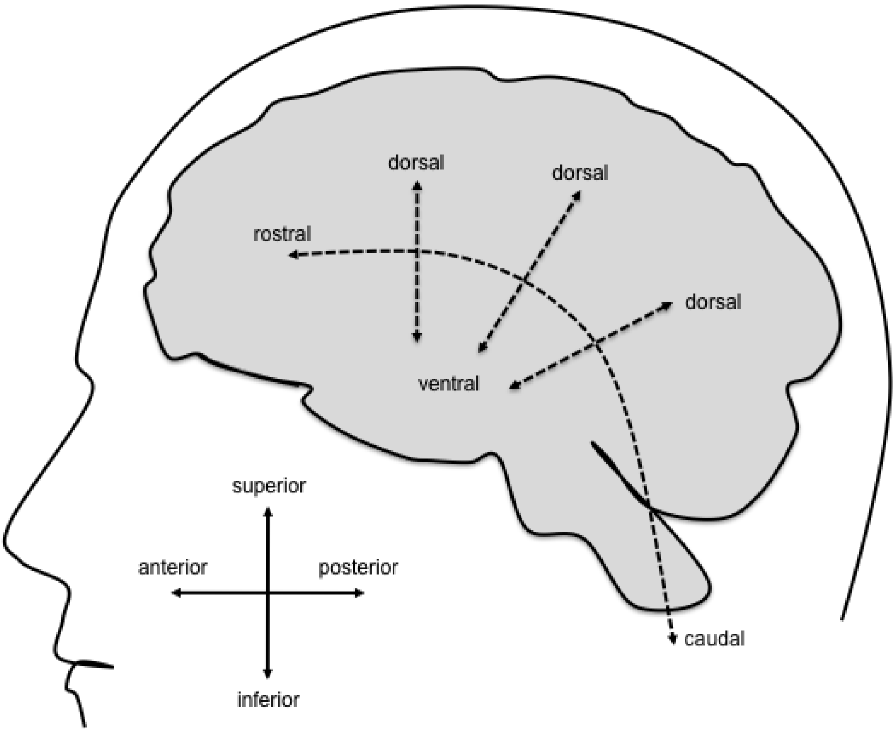


Figure 3-1. Coordinate systems that are used in cognitive neuroscience include one that has been adapted from animal studies, where the rostral-caudal axis is curved, in contrast to the anterior-posterior coordinate axis that is a straight one. See text for details.

To add to the confusion, there is yet another coordinate system in neuroimaging that communicates the directions along which the images slicing through the three-dimensional brain volume are oriented. Sagittal refers to images that are taken orthogonal to the left-right axis of the brain (*i.e.*, looking at the brain from the side). Axial (or transverse) refers to images that are taken orthogonal to the superior-inferior axis and, finally, coronal implies that the image has been taken orthogonally to the anterior-posterior axis. The left and right side of the images are often marked due to there being two conventions: in the radiological convention, left side of the brain on axial / coronal images is located on the right side of the image and *vice versa* (*i.e.*, as if you were looking at the image from the bottom rather than from the top of the head in axial images). In the neurological convention, the left side of the brain is on the left of the image, and the right side of the brain is on the right of the image.

Knowing these coordinate systems is useful when reading through cognitive neuroscience literature, and helps one to understand where things are located with respect to one another. In the following, an introduction to major divisions of brain anatomy is first described, followed by more detailed description of the structures and brain areas considered as most relevant for perceptual and cognitive functions. Naturally, more detailed descriptions of how the anatomical areas are associated with specific cognitive functions are found in the subsequent chapters of this book.

3.2 Central and peripheral nervous systems

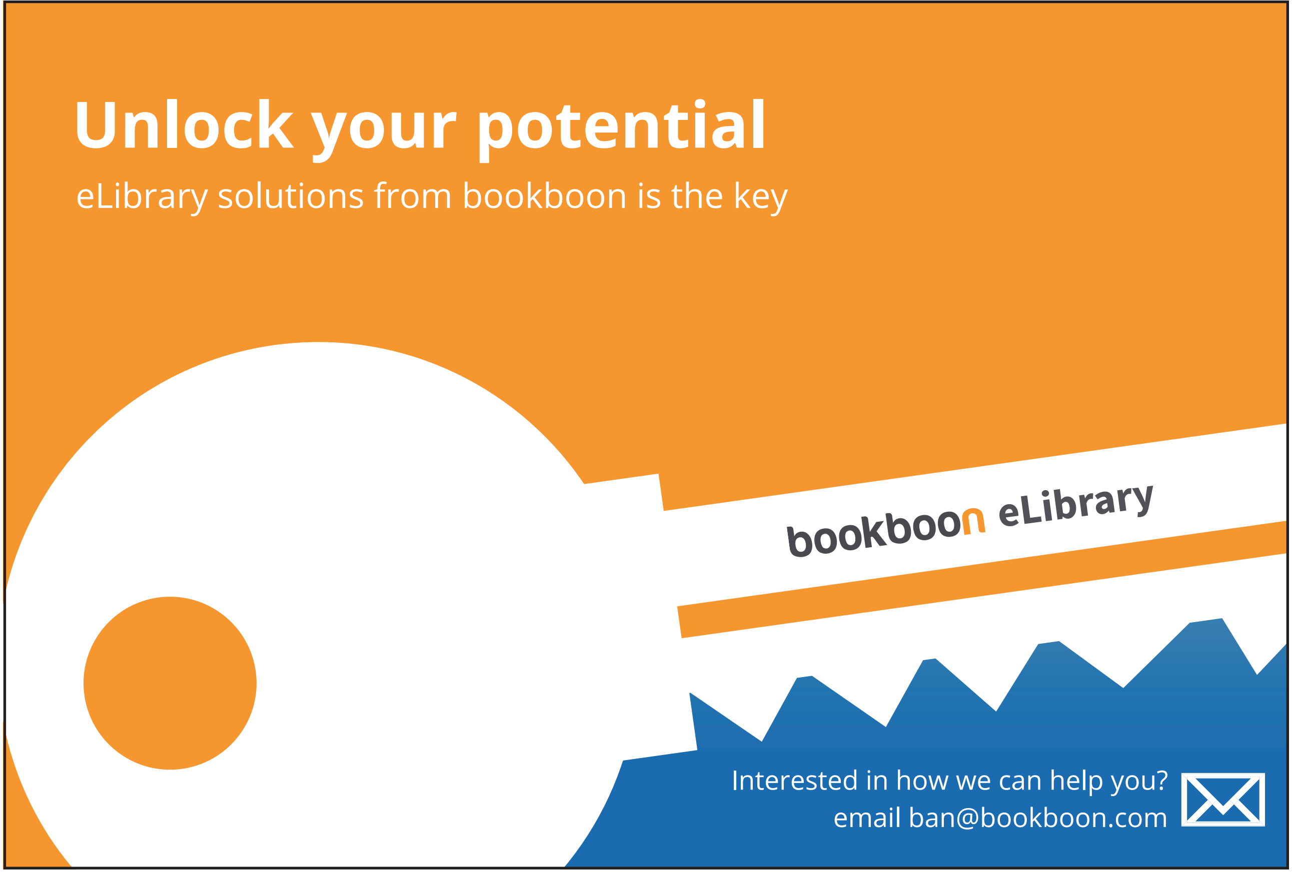
The brain is only a part of a larger nervous system that extends to each part of the body. The nervous system can be divided into central nervous system (CNS), which consists of the brain and the spinal cord, and peripheral nervous system that consists of the neurons of sensory systems, such as cells that react when something touches one's skin, and specific types of cells in one's ears that are sensitive to sounds. In contrast to the neurons that bring information from the senses to the central nervous system, the motor neurons belong to the CNS as the cell bodies of these neurons are within the CNS even though they send inputs all the way down to the most distal parts of the body thus making the muscles move.

Neurons of the so-called autonomic nervous system also belong to the peripheral nervous system (for description of neurons, see Chapter 4). The autonomic nervous system (for a review, see (Shields, 1993)) is divided into sympathetic and parasympathetic systems. The sympathetic nervous system has an excitatory effect on the body, producing a so-called "flight-or-flight" or stress response; the heart beats faster and stronger, lungs work harder to get more oxygen, pupils of the eye dilate, blood pressure increases, and the function of the digestive tract is slowed down. Activation of the parasympathetic nervous system exerts opposite, calming, effects on the body.

3.3 Gross anatomical divisions of the brain

The brain can be divided into evolutionarily older and younger structures. The evolutionarily older structures are found in the medial-inferior parts of the brain that are best seen in a mid-sagittal view of the brain shown in **Figure 3-2**. Extending from the spinal cord, at the base of the brain, is the so-called brainstem that encompasses several distinct anatomical structures. Starting from the most caudal part of the brainstem, connecting the brain stem to the spinal cord, medulla oblongata regulates vital body functions such as heart rate and breathing. Superior to medulla, the pons relays sensory and motor inputs but, additionally, is vital for the regulation of arousal, including sleep.

Superior to the pons, a structure named tectum contains inferior and superior colliculi that are highly important structures for early processing (and also in the case of superior colliculus fusion of) auditory and visual information, respectively. Another structure that is located superior to the pons is the tegmentum. The ventral tegmental area has been intensively



studied as the source of specific neurochemical transmitter substance dopamine that has been associated with reward circuitry of the brain that is important for motivation and is also implicated in development of drug addictions.

Cerebellum is located posterior to the pons. The cerebellum is a structure that has been implicated in coordination/timing of well-learned motor sequences, such as a golf swing executed by an experienced golfer. Damage to the cerebellum results in deficits in fine motor coordination, balance, and motor learning. The cerebellum gets inputs from sensory and motor cortical areas of the brain and sends output through the so-called deep nuclei to thalamus and cortical (especially premotor and motor) areas.

Recently, there has been increasing recognition that the cerebellum, in addition to its role in motor coordination, is important for cognitive functions such as selective attention and emotions (Schmahmann, 2010). The posterior lobe of the hippocampus is especially implicated in other than motor functions. Thalamus, on the other hand, is a very important structure through which most of the sensory inputs travel to cortex (as well as being a major target of cortical descending inputs) and hypothalamus regulates hormonal secretion *via* the pituitary.

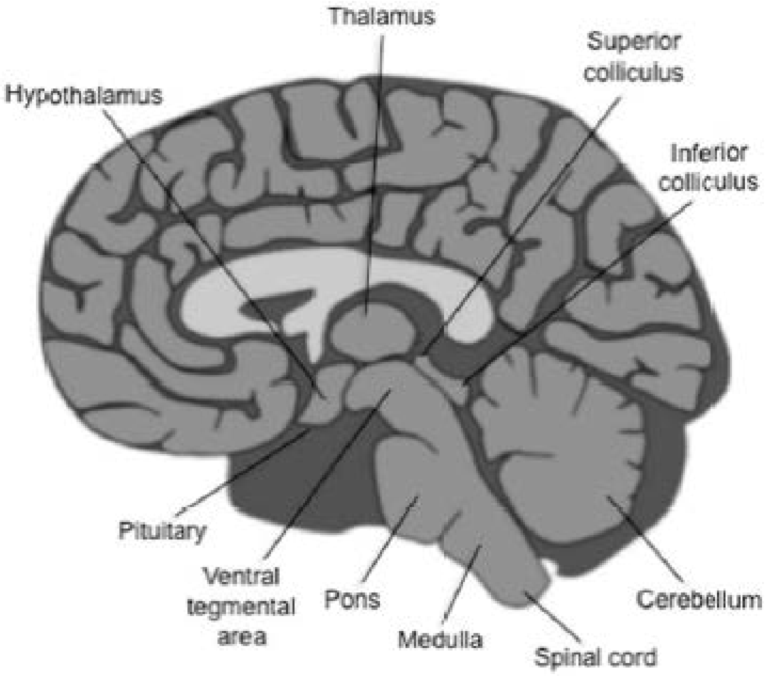


Figure 3-2. Evolutionarily older brain structures indicated in a mid-sagittal view of the brain. See text for details.

To summarize, the evolutionarily older structures that underlie the cerebral hemispheres are mainly responsible for regulating vital body functions, arousal, and relaying ascending inputs from the senses (and also descending motor commands). If humans only had these structures, we would however, lack what makes us human; the ability to reason, remember, plan ahead, feel for others, be social, possess sense of humor etc. All these are made possible by the cerebrum that is the part of the brain that has expanded the most during evolution.

3.4 The cerebrum

The cerebrum is clearly the most interesting part of the brain from the point of view of cognitive neuroscience. The cerebrum consists of highly symmetrical left and right hemispheres that seem to be functionally specialized to some degree, for instance, the left hemisphere appears to be more relevant for speech and language functions whereas the right hemisphere is more important for spatial processing and music. Each hemisphere can be further divided into four lobes: frontal, parietal, temporal, and occipital, based on the skull bones under which they reside (see **Figure 3-3**). Interestingly, as will be described later in this book, these different lobes seem to be responsible for distinct types of information processing.

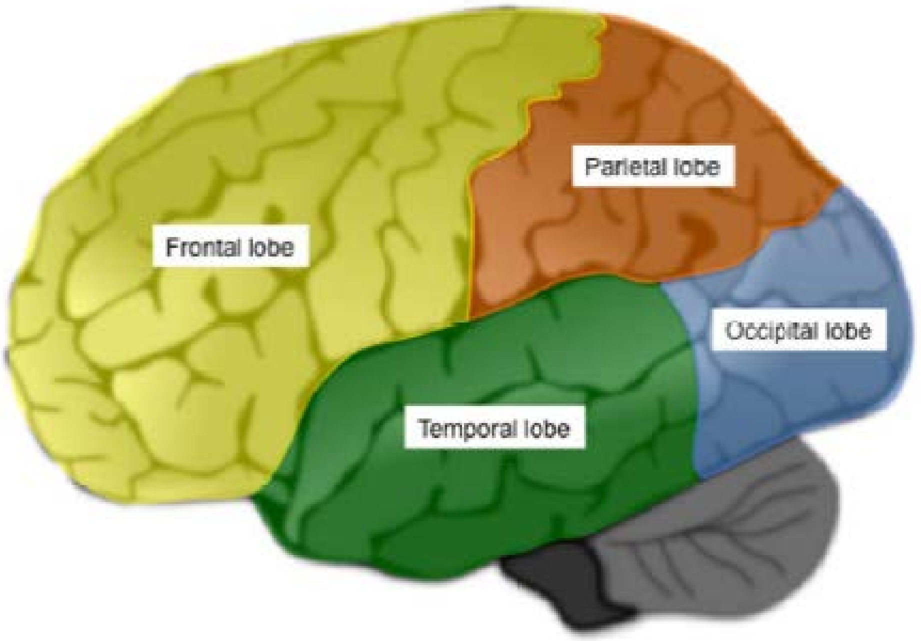


Figure 3-3. Shown are the frontal, parietal, temporal and occipital lobes of the left cerebral hemisphere. See text for details.

Gray and white matter

The cerebral hemispheres are made up of gray and white matter (see **Figure 3-4**). The white matter consists of so-called axons of neurons that carry neural impulses over longer distances (see Chapter 4). Fatty substance called myelin envelopes the axons, making the signals go up to 10 times faster, and giving the characteristic white color to the nerve bundles. The gray matter that makes up the cortex (*i.e.*, the surface of the brain) consists of six layers of neurons with unmyelinated fibers. The thickness of the cortex varies, depending on the part of the brain, from 2 to 4 mm, and is in most parts of the cortex about 3 mm.

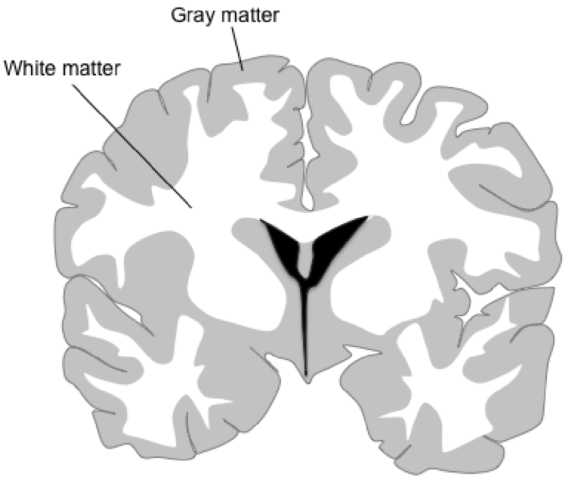


Figure 3-4. As can be seen in this mid-coronal section of the brain, the cerebral hemispheres are composed of gray and white matter.

Brodmann areas

One way that the cortex has been parcellated is based on fine-grained anatomical differences in the composition of the cortical layers: their thickness and types of neurons that are found in each layer. By mapping these differences, German neurologist Korbinian Brodmann was able to define what are still today known as Brodmann areas (Brodmann, 1909). The Brodmann areas are depicted in **Figure 3-5** below. Brodmann-area based labeling of relevant brain regions is still used in modern cognitive neuroscience, even though there are more advanced alternative methods increasingly available (Van Essen et al., 1998).

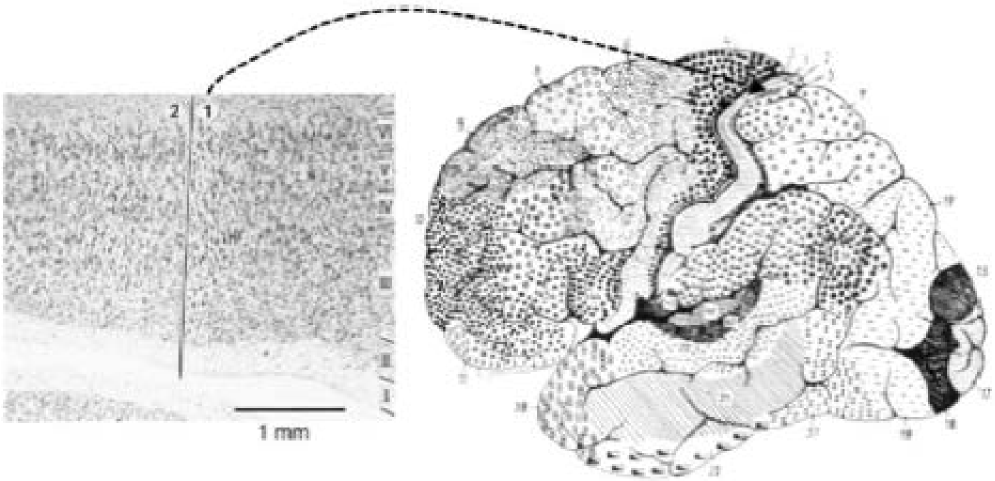


Figure 3-5: Division of cortex into Brodmann areas based on cytoarchitectonic differences between different cortical areas. Shown on the left side is an example of fine cytoarchitectonic differences between adjacent cortical areas, in this case between Brodmann areas 1 and 2. The cortical layers are indicated with roman numerals. On the right, the Brodmann areas are shown on the lateral surface of the brain. Adpated from (left) (Geyer et al., 1999) and (right) (Brodmann, 1909).

3.5 The cerebral cortex is folded into gyri and sulci

The lobes of the cerebral hemispheres consist of gyri and sulci (*i.e.*, ridges and furrows) that are made up by folding of the cortex that make it possible "to pack" more cortex to the limited space available inside the cranial cavity. In fact, the cortex surface measures up to 2000 cm^2. Specific sulci (also in some cases called fissures) form the anatomical borders between the lobes (see **Figure 3-6**). Frontal and parietal lobes are separated by the central sulcus, anterior and posterior to which are the precentral and postcentral gyri, respectively. The Sylvian fissure, on the other hand, forms the border between temporal and frontal, as well as posterior temporal and inferior parietal cortex.

Gyri and sulci of the temporal lobe

The gyri and sulci on the lateral aspect of the temporal lobe (see **Figure 3-6**) follow a relatively straightforward nomenclature. Specifically, the gyri and sulci are named based on the lobe that they are located in and their relative position within the lobe in the superior-inferior dimension. The gyrus that is most superior on the lateral aspect of the temporal lobe, running along the anterior-posterior axis, is called superior temporal gyrus, the one inferior to it is the middle temporal gyrus, and the most inferior gyrus of the temporal lobe is called the inferior temporal gyrus. The sulcus between the superior and middle temporal gyri is called the superior temporal sulcus, and the one separating the middle and inferior temporal gyri has been named inferior temporal sulcus. The superior aspect of the temporal lobe is especially important for processing of auditory information including speech sounds.

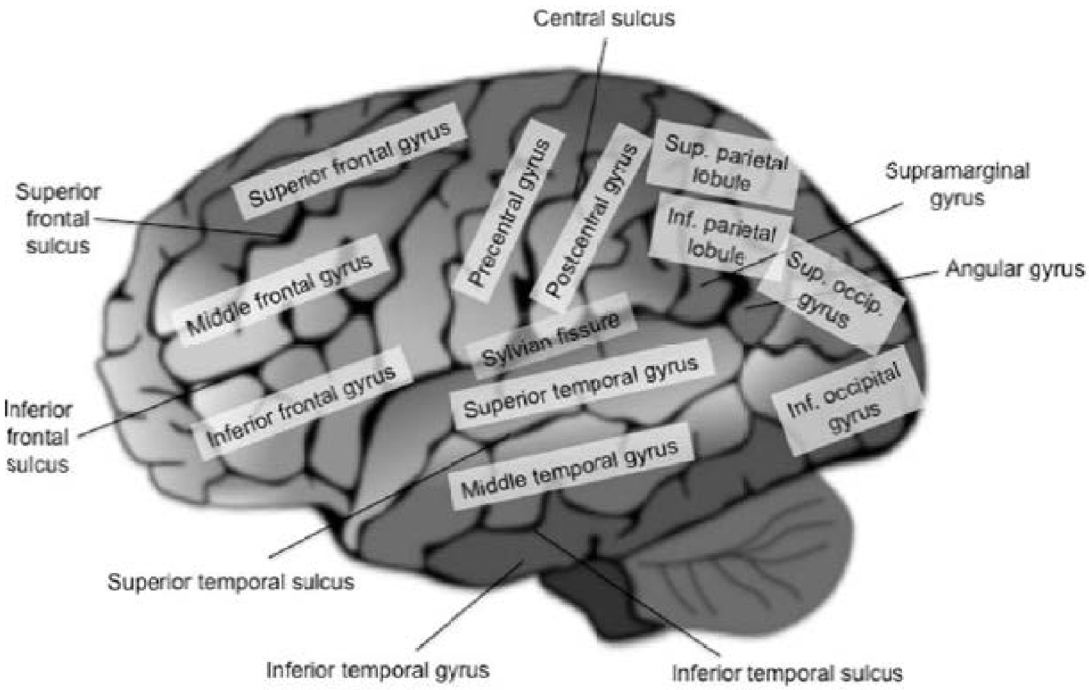


Figure 3-6. Shown are the gyri and sulci of the cortex on the lateral view of the left cerebral hemisphere. The right hemisphere is highly similar with the left hemisphere. See text for details.

When one lifts the frontal and parietal lobes off of the superior aspect of the temporal lobe and looks into the Sylvian fissure, there are several transverse gyri and sulci that run orthogonal to the Sylvian fissure. The most posterior of these is the planum temporale that has been associated with processing of the locations of sounds, and anterior to it is the Heschl's gyrus, the most medial two-thirds of which form the primary auditory cortex where most of the inputs from the subcortical nuclei of the auditory system first arrive. Anterior to these can be found the Heschl's sulcus, insula (that extends from temporal lobe to the inferior aspect of the frontal lobe), and planum polare.

On the inferior aspect of the temporal lobe, one finds, in the junction between the inferior temporal and inferior occipital lobes, the occipito-temporal fusiform gyrus that contains cells that are sensitive to seeing faces. Traversing into the medial parts of the temporal lobe, there is entorhinal cortex that runs along the anterior-posterior direction and medial-superior to it the hippocampus that is a structure highly relevant for encoding of memories (as will be described in more detail in Chapter 7).

Gyri and sulci of the frontal lobe

In the frontal lobe, the gyri that run along the anterior-posterior direction are called the superior, middle, and inferior frontal gyri. The two sulci that separate middle frontal gyrus from superior and inferior frontal gyri are called superior frontal sulcus and inferior frontal sulcus, respectively. The frontal cortex is especially important for motor coordination but also for higher-order cognitive and emotional functions such as planning of goal directed actions and social cognition (*i.e.*, how we perceive the intentions of others and adjust our own behavior accordingly; these functions are described in Chapters 10 and 12). In fact, the frontal lobe is the brain structure that has grown the most in humans compared with other species (see **Figure 3**-7).

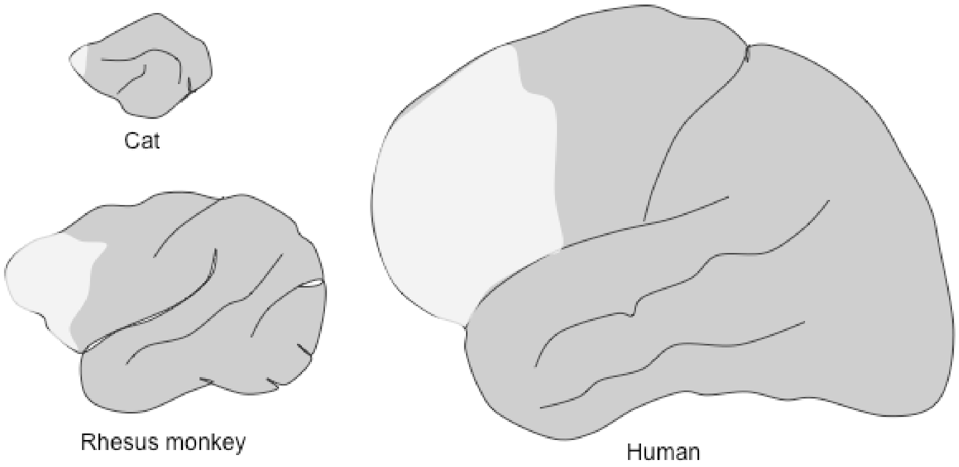
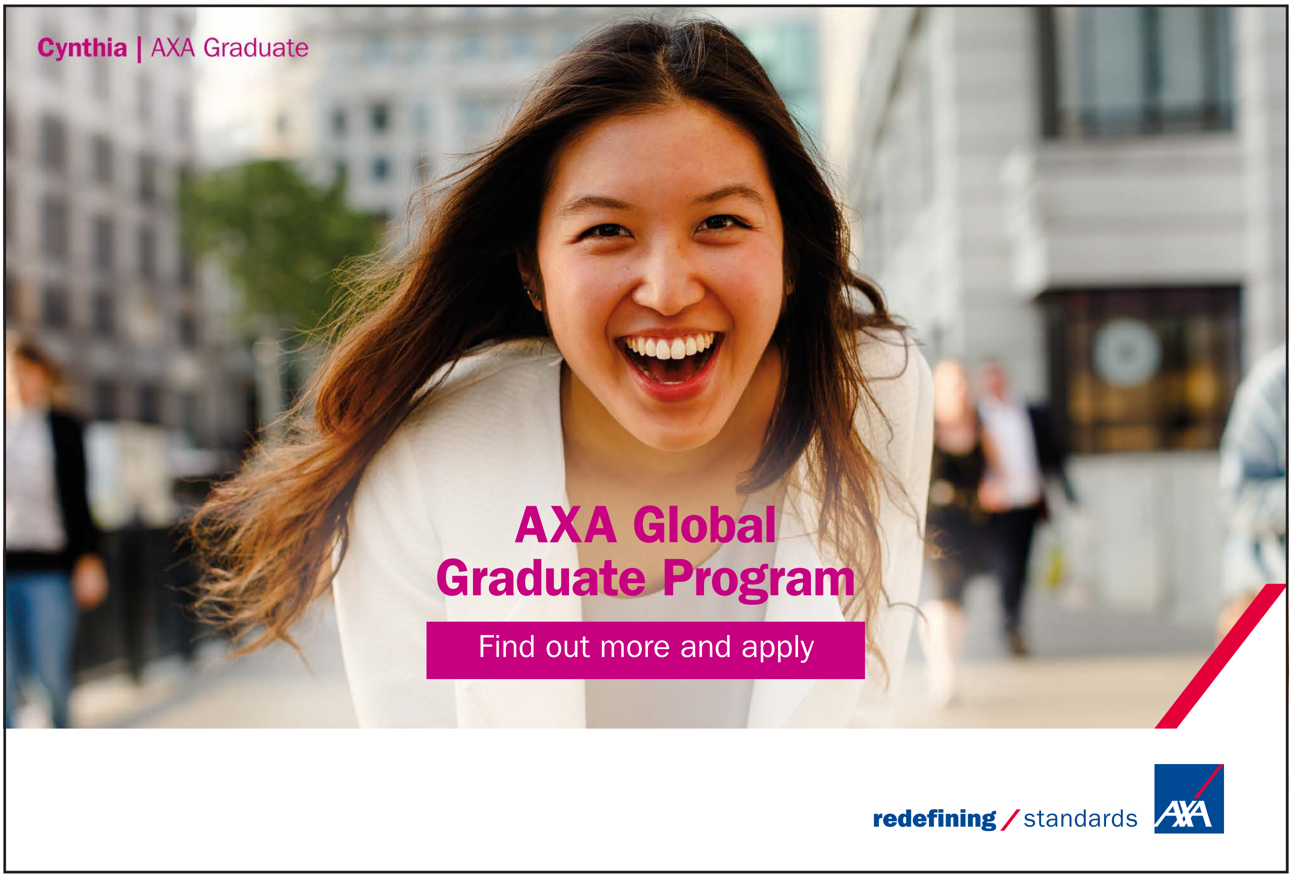


Figure 3-7. Shown is a comparison of prefrontal cortex size between cat, rhesus monkey, and human. The prefrontal cortex is indicated with lighter shade of gray. As can be seen, the proportional size of the prefrontal cortex is most expanded in humans.

In the frontal lobe there are also gyri and sulci that do not follow the superior-inferior nomenclature. The inferior frontal gyrus is composed of two smaller gyri with the more posterior one called pars triangularis and the anterior one pars opercularis. Together, these gyri have been, especially on the left hemisphere, associated with speech production but also with speech perception and are also referred to as the Broca's area. The anterior tip of the frontal lobe is called the frontal pole, but there is no definite agreement as to how extensive an area the frontal pole encompasses. Adding to confusion, approximately the area of middle frontal gyrus is often also called the dorsolateral prefrontal cortex, an area that is important for attentional/volitional functions (see Chapter 6).



Peeking on the other (medial) side of the hemisphere (see **Figure 3-8**), one can find, buried in the fissure between the hemispheres, the so-called supplementary motor area in the superior aspect of the medial frontal lobe (occupying the anterior portion of paracentral lobule) that is highly important for planning of motor sequences (Roland et al., 1980). Other notable brain areas include the ventromedial prefrontal cortex that plays an important role in social cognition (see Chapter 12), and residing posterior to this area in the ventral-medial-basal aspects of the frontal lobe is the so-called orbitofrontal cortex that is important for regulation of emotions (see Chapter 11).

The cingulate gyrus that extends into the medial part of the frontal lobe is often not considered to be a part of the neocortex, as it differs microanatomically from other cortical areas, but rather as a part of the so-called limbic system, which will be described later in this chapter. The anterior cingulate gyrus has been associated with information processing that takes place when information from separate sources or between desired responses are in conflict with each other (Carter et al., 1998).

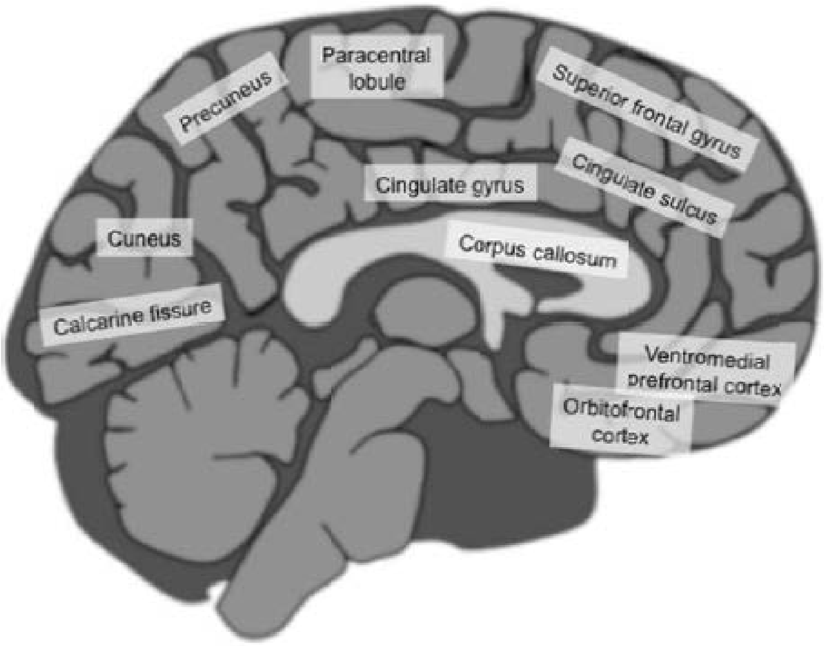


Figure 3-8. Medial view of the left cerebral hemisphere is shown with anatomical structures labeled. See text for details.

The posterior aspect of the frontal lobe contains the primary motor cortex and on the other side of the central sulcus that separates frontal and parietal lobes there is the primary somatosensory (*i.e.*, senses of pressure/touch, heat/cold, and pain) cortex. These precentral and postcentral gyri contain topographic representations of the body that are called the motor and somatosensory homunculi, respectively (Penfield and Jasper, 1954) (see **Figure 3-9**). What is peculiar about these homunculi is that they are distorted, with some body areas disproportionately represented. This corresponds to the degree of fine motor coordination and tactile resolution of, and thus increased number of neurons that innervate, these body parts.

The tactile resolution of different body parts can be tested using the so-called two-point discrimination ability test. You can try this yourself by touching your skin simultaneously by tips of two pencils placed adjacently. When touching your finger, the pencil tips can be relatively close to one another and still you can feel that they are separate, whereas if you touch your leg, the pencil tips need to be separated by a much larger distance in order to feel them as separate.

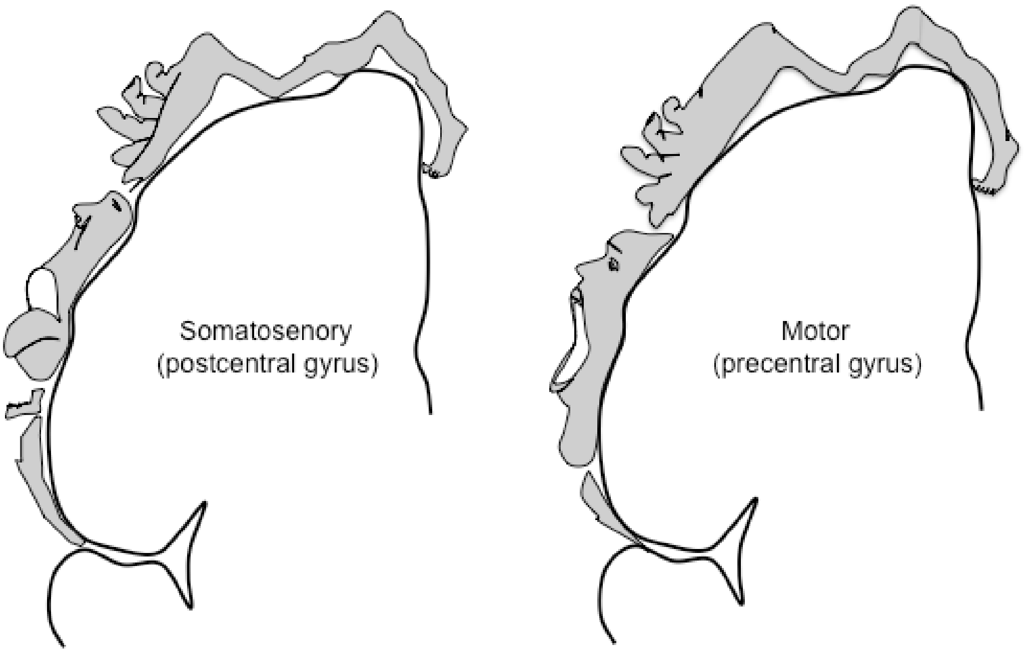


Figure 3-9. Somatosensory (on the left) and motor (on the right) homunculi of the left hemisphere post- and pre-central gyri are here shown on a coronal section. Note the disproportionately large representation areas of certain body parts with high degree of motor coordination and tactile sensitivity such as fingers.

Anatomical structures of the parietal lobe

Besides the primary somatosensory cortex located in the post-central gyrus at the anterior border of the parietal lobe that contains the somatosensory homunculus shown, there are so-called inferior parietal and superior parietal lobules on the lateral-posterior aspect of the parietal cortex (see **Figure 3-6**) that have been associated with processing of spatial information (*i.e.*, where things are) as well as perceiving and executing actions in space (for a review, see (Andersen et al., 1997)). In the junction between superior-posterior temporal lobe and inferior parietal lobe, there are gyri that have been named supramarginal gyrus and angular gyrus. In the medial aspect of the parietal lobe, one finds the posterior aspect of the paracentral lobule and cingulate gyrus, as well as precuneous.

Anatomical areas of the occipital lobe

The occipital lobe in the posterior part of the brain is especially important for processing of visual information that is transmitted there from the eyes *via* the thalamus. The primary visual cortex is located on the medial aspect of the occipital cortex, within the so-called calcarine fissure (see **Figure 3-8** above). Cuneus is located superior to the calcarine fissure on the medial aspect of the occipital lobe, and the lingual gyrus is inferior to the calcarine fissure. On the lateral side of the occipital lobe, the lateral occipital sulcus separates the superior and inferior occipital gyri. In the lateral-inferior aspect of the occipital lobe, there is a functionally defined anatomical area called the lateral occipital complex that has been reported to hold distributed representations of different perceptual objects (Ishai et al., 1999) (distributed representation refers to spatial activity patterns giving rise to specific representations, as will be described in more detail in Chapter 7).

3.6 Beneath the cerebral hemispheres: the limbic system and basal ganglia

The so-called limbic system is a collection of brain structures that are located just beneath the cerebral hemispheres (see **Figure 3-10**) -- the term limbic comes from Latin "*limbus*" that means "border". Rather than being a unitary functional system (as was at one time hypothesized) the brain structures that belong to the limbic system serve quite distinct functions (for a review see (Isaacson, 2003)). The cingulate gyrus is the most superior structure of the limbic system and together with parahippocampal gyrus and dentate gyrus forms the so-called limbic lobe. The corpus callosum (shown in **Figure 3-8**) that resides underneath the cingulate gyrus consists of fibers that connect the two hemispheres. Anterior and posterior commisures also serve this function.

Hippocampus is a structure of the limbic system that is highly important for memory functions (see Chapter 7). Located at the anterior tip of the hippocampus (bilaterally as is usually the case in brain anatomy) is the amygdala. The amygdala has been observed to play a critical role in emotional processing (see Chapter 11), which is the function that the limbic system once was thought to underlie as a whole. Other structures found in the limbic system include fornix, septal nuclei and the mamillary body.

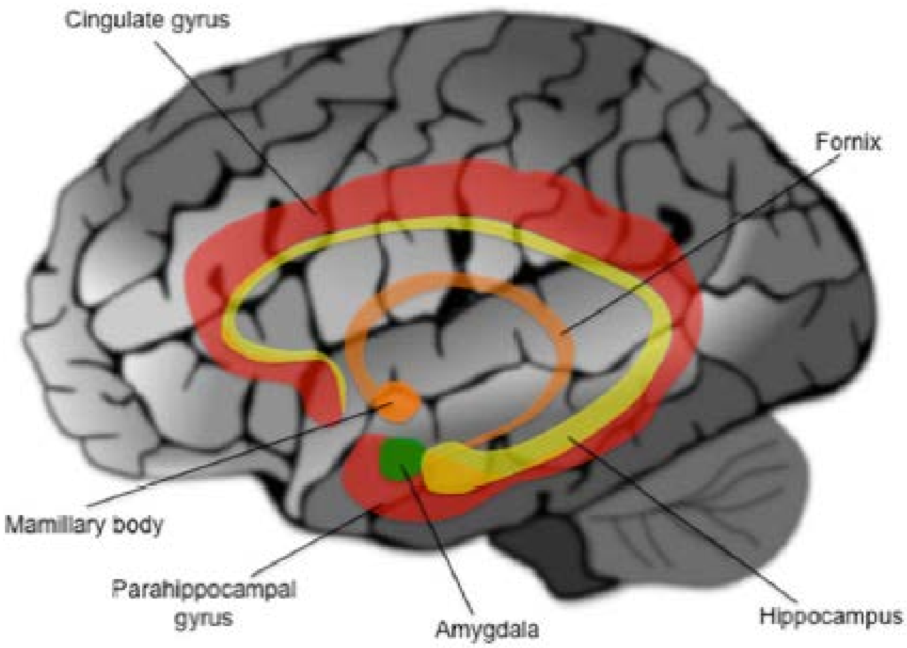


Figure 3-10. Shown are the anatomical structures of the limbic system that are located deep inside the brain. Here, they are projected to the cortical surface for visualization purposes.

Basal ganglia are another group of subcortical structures residing underneath the frontal lobes. Basal ganglia are a highly important group of nuclei for movement control, initiation of movements/actions, and also for higher cognitive functions. The basal ganglia are connected with other brain structures, including prefrontal cortical areas and the thalamus. Together with these structures, the basal ganglia form a functional network that is responsible for selection of an action out of many possible alternative actions. The structure and function of basal ganglia are described in more detail in Chapter 10 of this book.

3.7 The ventricular system

While the ventricular system is naturally not involved in cognitive processing since it consists of cavities filled with cerebrospinal fluid, it is good to acquaint oneself with the ventricular system as the ventricles are prominent landmarks in the brain. The ventricular system, shown in **Figure 3-11**, consists of the central canal that runs from the spinal cord up to the base of the brain where it fuses into the cerebral aqueduct. The so-called third ventricle and, bilaterally, left and right lateral ventricles reside superior to the cerebral aqueduct. Abnormal enlargement of ventricle size is a robust sign of pathology that can be seen with neuroimaging methods such as magnetic resonance imaging described in Chapter 2.

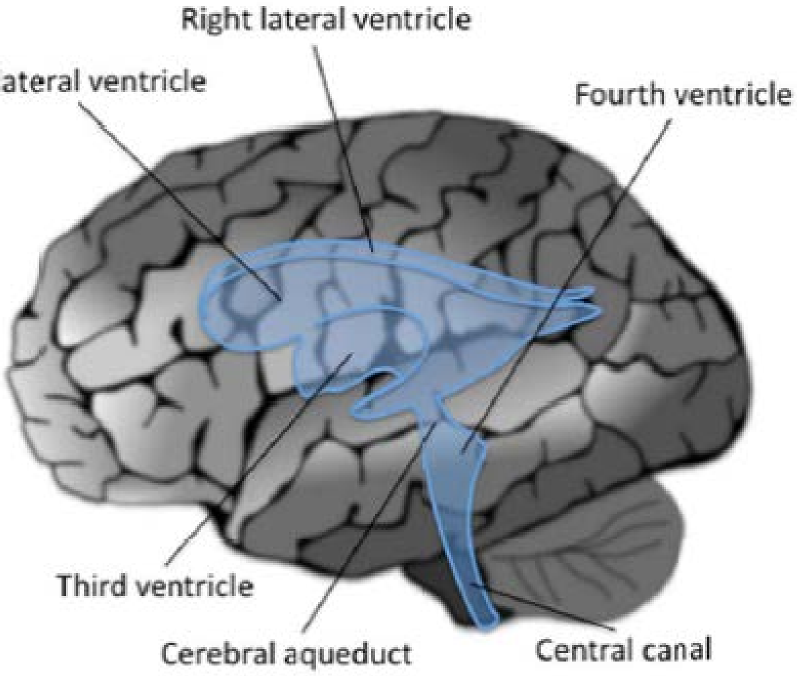


Figure 3-11. Shown is the ventricular system of the brain. The ventricles offer prominent anatomical landmarks when browsing through three-dimensional brain volumes obtained using, for example, magnetic resonance imaging. The ventricles are located deeper inside the brain and are here projected to the cortical surface for visualization purposes.

3.8 Connectivity of the brain

Even though most of the attention in cognitive neuroscience has been focused on studying the role of gray matter brain areas / structures in perceptual and cognitive functions, it is increasingly recognized that brain connectivity (*i.e.*, the tracts that run in the white matter) plays an important role in cognition. This is due to the fact that even though specific functions can be associated with single brain areas, a closer inspection reveals that it is usually a network of brain areas that, through intricate interactions and modulation of other brain areas, underlies cognitive processes. There are different types of connections in the brain. From the sensory organs, there are ascending fibers that traverse through specific nuclei in brain stem and thalamus to cortex *via* multiple synapses. On the other hand, there are an order of magnitude more descending fibers from the cortex to thalamus and the other subcortical structures than there are ascending fibers, making it possible for cortex to modulate processing in the lower order areas that is relevant for selective attention (see Chapter 6).

The thalamus

The thalamus, the left side of which is pictured in **Figure 3-12**, is the subcortical structure that is most densely connected with the cortex. In fact, the thalamus is composed of multiple nuclei and it is often said to be a miniature of the respective cerebral hemisphere, as posterior parts of the thalamus are connected to posterior parts of the cerebral cortex and *vice versa*. In addition to the sensory nuclei of the thalamus, the most notable of which are perhaps the middle geniculate body (MGB) that connects the thalamus to the auditory cortex and lateral geniculate nucleus (LGN) that connects the thalamus to the visual cortex, there are nuclei that help regulate arousal and support higher cognitive functions. The thalamus also connects to the motor and prefrontal cortices. Notably, the number of connections that descend from cortex to thalamus is higher than the number of ascending connections. The role of the descending connections is presumed to be a regulating one that allows top-down filtering of sensory information (see Chapter 6).

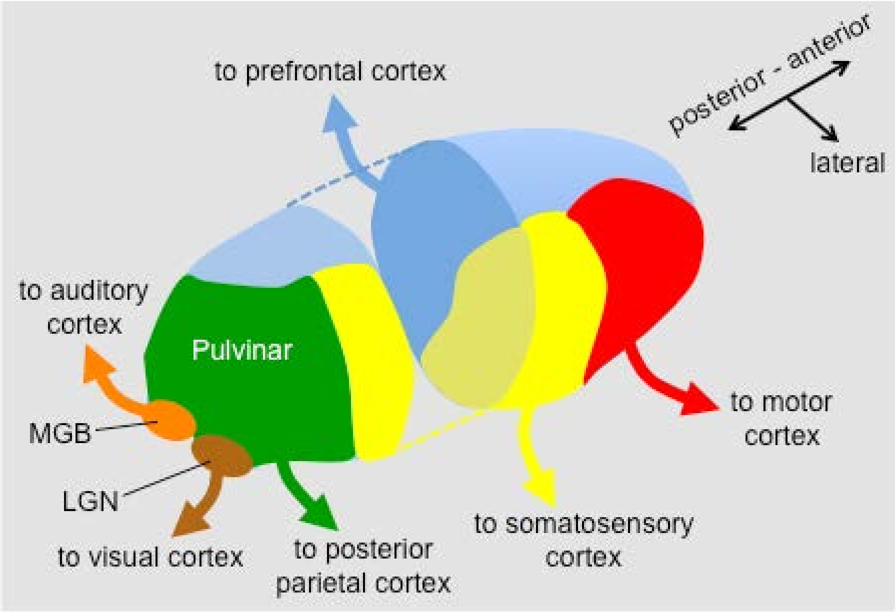
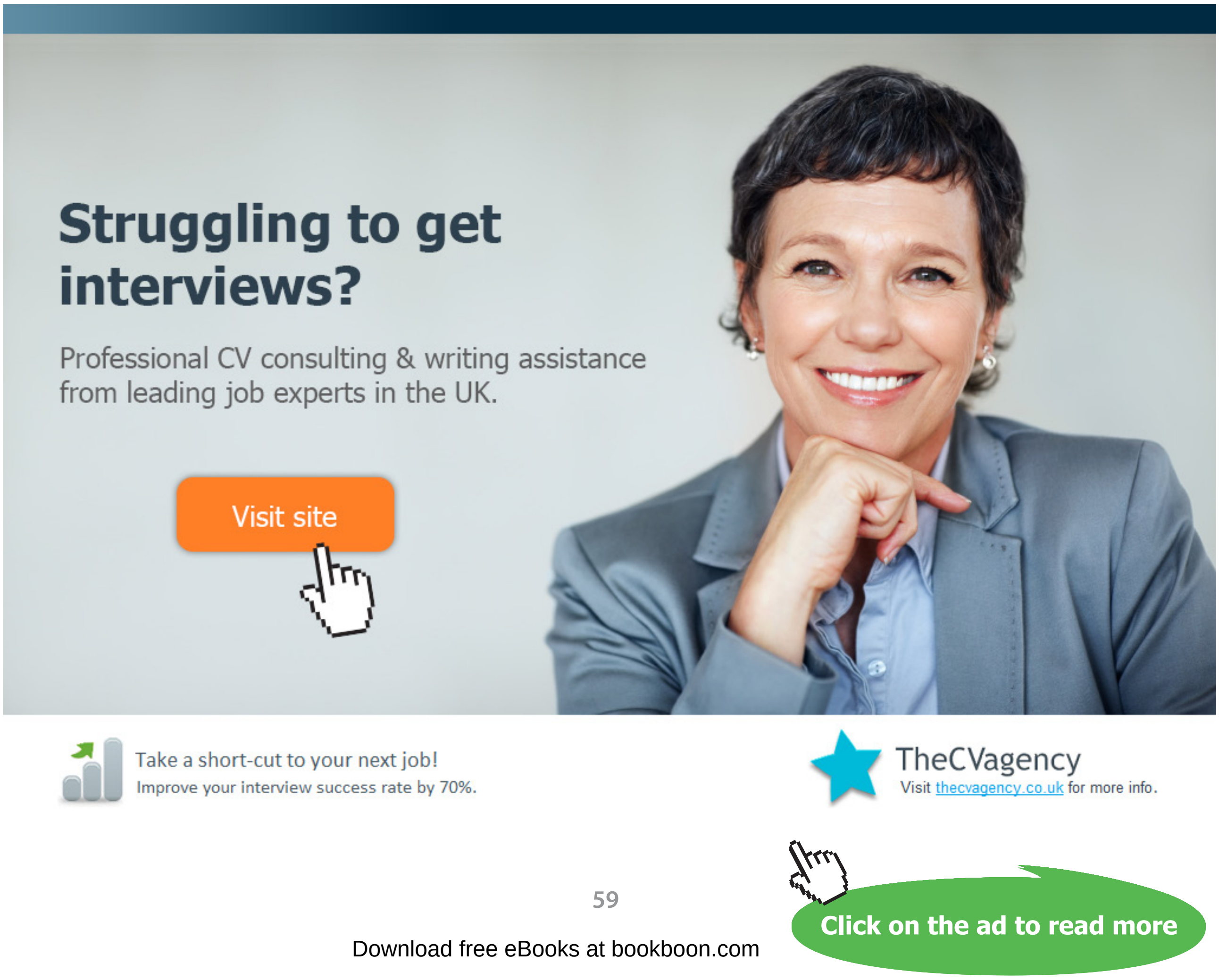


Figure 3-12. Shown are the main nuclei and ascending connections of the left thalamus. Here, the thalamus is sliced in two along the anterior-posterior dimension. As can be seen in this simplified illustration of the thalamus, there are a number of nuclei contained in the thalamus and the posterior parts of the thalamus connect to posterior cortical areas and anterior parts of the thalamus connect to the anterior parts of the cerebral cortex. Descending connections (not shown here) follow the same connectivity pattern and are higher in number than the ascending connections.



The corpus callosum connects the two hemispheres and there are additional inter-hemispheric connections that run through the posterior and anterior commissures. In addition to the inter-hemispheric and ascending-descending connections, there are so-called U-fibers that connect adjacent cortical areas to one another. The density or number of connections does not necessarily dictate the strength of the connections. As will be described in the next Chapter, the type of chemical neurotransmitter and the type of receptor that the neurotransmitter binds to determine whether the connection is an excitatory or inhibitory one.

3.9 Concluding remarks

Adequate knowledge of brain anatomy is a prerequisite for in-depth understanding of cognitive neuroscience. While the anatomical nomenclature with multiple synonyms can be initially confusing, and the shapes and locations of the various structures might be difficult to grasp given that they are located in the oddly shaped three-dimensional cranial cavity, it does not take long to become accustomed with the few fundamentals of brain anatomy. Knowing the different coordinate systems and major divisions of the brain already helps one when reading through cognitive neuroscience literature. It is a good practice to stop and check where the different brain areas are that are mentioned when reading through the chapters that will follow in this book. After all it is not the anatomy in by itself that is remarkable to a student of cognitive neuroscience, but rather the roles played by the various structures (and by the interactions of networks of anatomical structures) in functions such as perception, memory, decision-making, emotions, social interactions, and consciousness.

3.10 References

ANDERSEN, R. A., SNYDER, L. H., BRADLEY, D. C. & XING, J. 1997. Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annual Review of Neuroscience,* 20**,** 303-330.

BARTLEY, A. J., JONES, D. W. & WEINBERGER, D. R. 1997. Genetic variability of human brain size and cortical gyral patterns. *Brain,* 120**,** 257-269.

BRODMANN, K. 1909. *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues,* Lepizig, Johann Ambrosius Barth.

CARTER, C. S., BRAVER, T. S., BARCH, D. M., BOTVINICK, M. M., NOLL, D. C. & COHEN, J. D. 1998. Anterior cingulate cortex, error detection, and online monitoring of performance. *Science,* 280**,** 747-749.

GREYER, S., SCHLEICHER, A., ZILLES, K. 1999. Areas 3a, 3b, and 1 of human primary somatosensory cortex. *Neuroimage*, 10, 63-83.

ISAACSON, R. L. 2003. *Limbic system*, eLS.

ISHAI, A., UNGERLEIDER, L. G., MARTIN, A., SCHOUTEN, J. L. & HAXBY, J. V. 1999. Distributed representation of objects in the human ventral visual pathways. *Proc Natl Acad Sci USA,* 96**,** 9379-9384.

PENFIELD, W. & JASPER, H. 1954. *Epilepsy and the functional anatomy of the human brain,* Oxford, Little, Brown & Co.

ROLAND, P. E., LARSEN, B., LASSEN, N. A. & SKINHOJ, E. 1980. Supplementary motor area and other cortical areas in organization of voluntary movements in man. *J Neurophysiol,* 43, 118-136.

SCHMAHMANN, J. D. 2010. The role of the cerebellum in cognition and emotion: personal reflections since 1982 on the dysmetria of thought hypothesis, and its historical evolution from theory to therapy. *Neuropsychol Rev,* 20, 236-260.

SHIELDS, R. W. J. 1993. Functional anatomy of the autonomic nervous system. *J Clin Neurophysiol,* 10, 2-13.

VAN ESSEN, D. C., DRURY, H. A., JOSHI, S. & MILLER, M. I. 1998. Functional and structural mapping of human cerebral cortex: solutions are in the surfaces. *Proc Natl Acad Sci USA,* 95, 788-795.

YANEZ, I. B., MUNOZ, A., CONTRERAS, J., GONZALEZ, J., RODRIGUEZ-VEIGA, E. & DEFELIPE, J. 2005. Double bouquet cell in the human cerebral cortex and a comparison with other mammals. *Journal of Comparative Neurology,* 486, 344-360.

4 Neurons

Like any other tissue in the human body, the various anatomical structures of the brain that were introduced in the previous chapter are composed of different types of cells. The majority of these cells are called glia cells. The main purpose of glia cells is to support the structure and metabolic functioning of the brain. There are different types of glia cells. Astrocytes help form the so-called blood brain barrier that keeps unwanted particles from entering the brain (Janzer and Raff, 1987). The blood brain barrier is also an obstacle when drugs targeting the central nervous system are developed, as the drug molecules need to pass through in order to exert their effects centrally.

Another type of glia, oligodendrocytes, speed up neuronal signaling by an order of magnitude by wrapping axons (that are specific parts of neurons, see below) into a myelin sheath (for a review, see (Hartline and Colman, 2007)). In the peripheral nervous system so-called Schwann cells serve the same purpose. However, unlike a single Schwann cell that sheaths only one axon, a single oligodendrocyte can sheath up to several tens of axons with myelin. Microglia are a third type of central nervous system glia that serve the purpose of cleaning the brain by removing dead tissue and filling up the resulting empty space (for a review, see (Hanisch and Kettenmann, 2007)).

In addition to glia cells, there are highly specialized cells in the brain called neurons. Even though neurons are 10 times less abundant than glia cells in the brain, it can be said that neurons are the elementary building blocks of cognition; neurons transmit and process information in the brain, thus making perception and cognition possible. Understanding the basic principles of how a single neuron functions, and how neurons form functional networks, is necessary for every student of cognitive neuroscience. In this chapter, the structure and functional principles of neurons are introduced.

4.1 Neurons are the elementary building blocks of cognition

While neurons come in many different shapes and forms (see **Figure 4-1** for examples of two distinct types of neurons: cerebellar Purkinje cells and cortical pyramidal cells), there are certain structures that are common to each neuron. A prototypical neuron is depicted in **Figure 4-2** below to illustrate these structures. As can be seen, the prototypical neuron is composed of dendrites that form a tree-like structure with far-reaching branches. The dendrite tree is connected to the soma of the neuron, and from the soma a single axon runs all the way down to terminal endings. These structures and the overall role they play in neural signaling are introduced in the following.

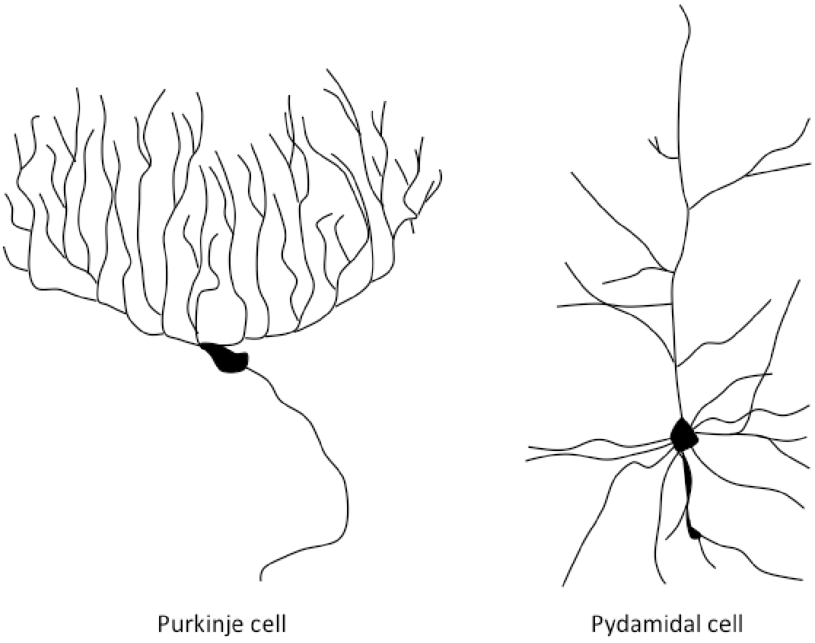


Figure 4-1: There are multiple different types of neurons in the brain. Here, shown are a cerebellar Purkinje cell on the left and a cortical pyramidal neuron on the right.

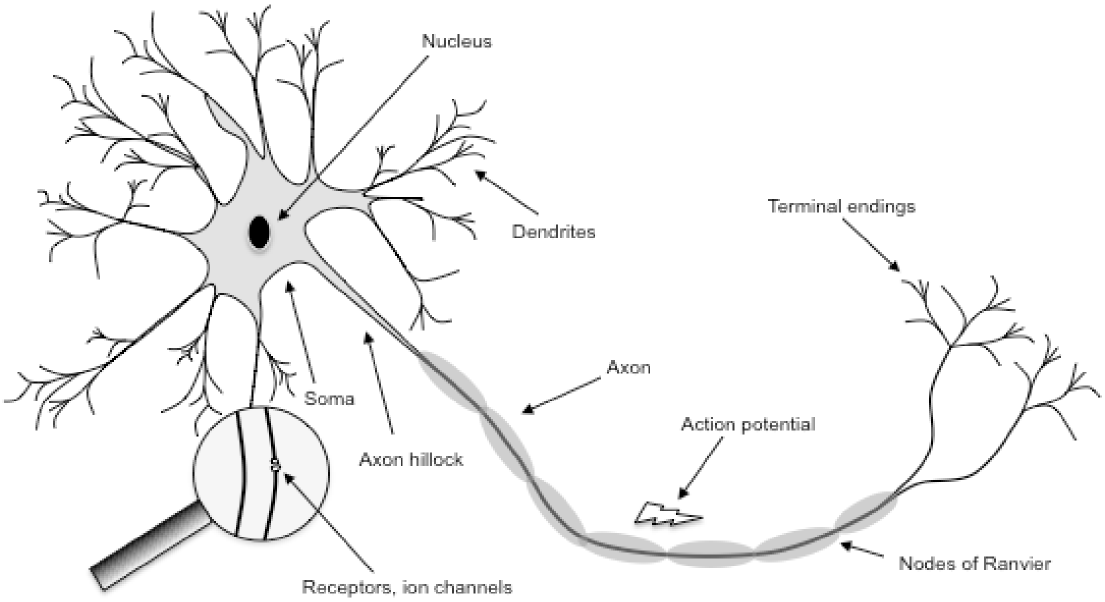


Figure 4-2: Shown are structures that are shared by all neurons. These structures include dendrites that receive inputs from other neurons by the means of chemical substances called neurotransmitters that are released by neurons and that bind to receptors on the dendrite cell membrane and result in opening/closing of specific ion channels on the cell membrane. The soma of the cell contains the nucleus and at the base of the soma there is the axon hillock that is the trigger zone of neuronal impulses. The axon carries the neural signal (called action potential) down to terminal endings that then results in release of neurotransmitters onto the dendrites of other neurons. Nodes of Ranvier are made up by glia cells that speeds up the propagation of the action potential up to tenfold.

Dendrites receive inputs from other neurons

Some neurons, such as the cerebellar Purkinje cells illustrated in **Figure 4-1** above, can have thousands of dendrites that receive inputs from other neurons. The inputs are conveyed to the recipient neuron at points of contact, called synapses, between terminal endings of the signaling neuron and the dendrites of the recipient neuron. The neuron that receives the signal is called the post-synaptic neuron and the neuron that sends the signal is called a presynaptic neuron. The signal is transmitted by release of specific chemical substances called neurotransmitters to the synaptic cleft by the presynpatic neuron. The neurotransmitter substance then binds to specific receptor proteins on the surface of the dendrite of the postsynaptic neuron.

When neurotransmitters bind to the receptors of the postsynaptic cell, ion channels are opened and/or closed depending on the type of the neurotransmitter and the type of the receptor, which affects the flow of ions to and from the cell. As the ions are charged either negatively or positively, their movement through the channels alters the so-called membrane potential (see below) of the neuron, an effect that (when it is a depolarizing one) can spread down the dendrite all the way to the axon hillock where the neuronal signal called action potential is then generated. A detailed description of the action potential, resting membrane potential, depolarization and hyperpolarization will be provided later in this chapter, after introducing the main structures of the neurons.

Soma is responsible for the metabolic function of the cell

The dendrites of the neuron connect to soma of the cell. The soma of a neuron is like that of any other cell. It contains the nucleus and is responsible for the metabolism of the cell. Without the soma, the neuron would not be able to sustain itself, as maintaining the resting membrane potential that enables neuronal signaling constantly requires energy. The depolarizing changes that travel down the dendrites are transferred through the membrane of the soma to the axon hillock that is located at the base of the soma. The axon hillock is the site where it is determined whether the neuron sends a neural impulse down its axon to dendrites of postsynaptic neurons.

Axon hillock, axon, and the action potential

At the base of the soma, there is the so-called axon hillock that is the site where the neural impulse (also known as an action potential) is initially generated. As will be described later in this chapter, there is a threshold in membrane potential that needs to be reached at the axon hillock by summation of depolarizing effects; once the threshold is reached, a single action potential is generated that travels down the length of the axon to terminal endings. The nodes of Ranvier that are made up by glia cells wrapping the axon on a myelin sheath speed up the propagation of the action potential.

Terminal endings

Once an action potential reaches the terminal endings, chemical neurotransmitter is released from the so-called vesicles (where transmitter is stored) into the synaptic cleft and onto the dendrites of the postsynaptic cells. As will be described later, the effect on the post-synaptic cell can be excitatory (*i.e.*, depolarizing) or inhibitory (*i.e.*, hyperpolarizing) depending on the type of the neurotransmitter and the type of the receptor. While a given pre-synaptic cell typically releases only a single type of neurotransmitter, there are multiple types of receptors in a given post-synaptic cell. There are also autoreceptors that the released neurotransmitter binds to in the terminal endings of the presynaptic cell. These autoreceptors provide feedback to the presynaptic neuron and help regulate the release of neurotransmitter. The mechanisms of neural signaling are described in detail in the following.

4.2 Mechanisms underlying neural signaling

In order to understand how neurons receive, integrate and pass on information, it is necessary to understand a few basic mechanisms that enable neural signaling. The first key mechanism is establishment/maintenance of the so-called resting membrane potential. The resting membrane potential allows for rapid local changes in membrane potential that occur during depolarization. And once the depolarization summates and spreads down to the axon hillock, an all-or-none action potential (also called neural impulse, spiking, or neural firing) is generated. These three key mechanisms will be described in the following. The principles of chemical messaging that takes place in the synapse will be described later in this chapter.

Resting membrane potential and the Na^+/K^+ pump

The ability of the neuron to signal with other neurons is dependent on continuous maintenance of the so-called resting membrane potential. This means that there is an electric potential difference between the extracellular and intracellular space of about 30–70 mV, with the intracellular being more negative than the extracellular aspect of the neural membrane. The resting membrane potential is maintained by the so-called Na^+/K^+ pump, which actively transports two K^+ ions from the extracellular to the intracellular space and three Na^+ ions from the intracellular to the extracellular space (see **Figure 4-3**). This creates a concentration gradient that forces K^+ out of the cell.

As the much larger, negatively charged, proteins/anions cannot pass through the ion channels of the neuronal cell membrane, a negative potential difference is generated (*i.e.*, the resting membrane potential is formed). On the other hand, Na^+ is attracted towards the intracellular space both by its concentration gradient and by its electrical gradient, but cannot pass through the cell membrane unless specific ion channels are opened. This makes it possible for the neural membrane to go very quickly from the negative resting membrane potential to a positive membrane potential (*i.e.*, when ion channels permeable to Na^+ do open). This phenomenon is called depolarization and it will be described next.

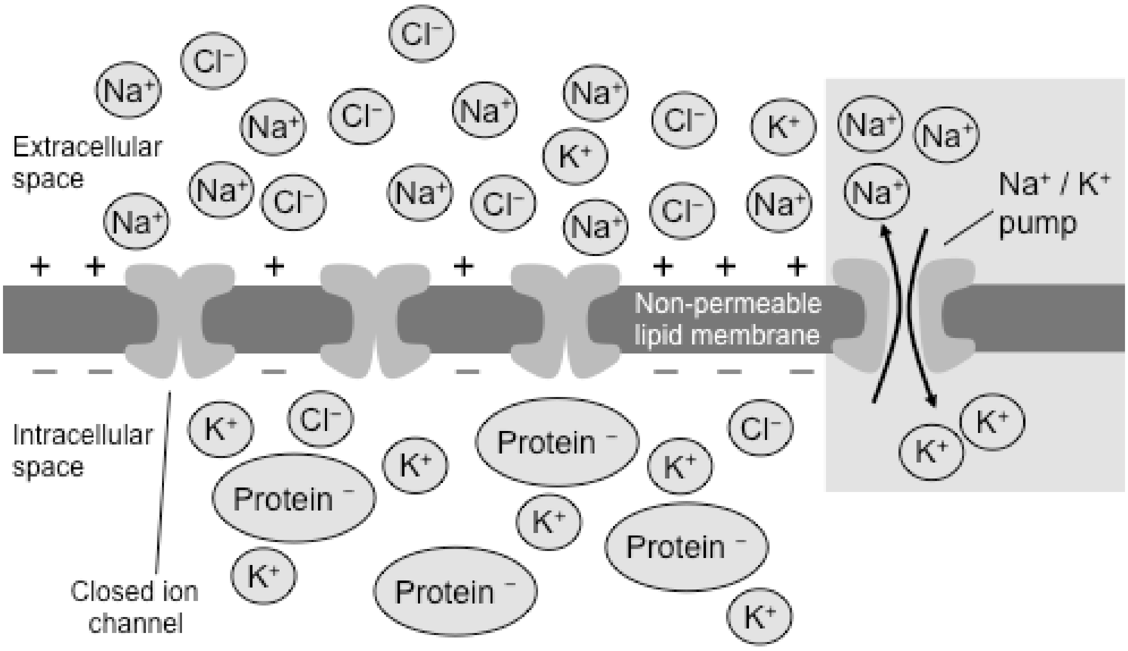


Figure 4-3. Schematic illustration showing how the resting membrane potential is formed. Shaded area on the right highlights the Na^+/K^+ pump. See text for details.

Depolarization and hyperpolarization

Depolarization occurs when positively charged ions (Na^+ and/or Ca^{2+}) enter the intracellular space through ion channels that are opened when neurotransmitters bind to receptors that regulate the ion channels (see **Figure 4-4**). Since there is the negative resting membrane potential, the positively charged ions (that are not pushed out of the cell by concentration gradient like the K^+ ions) are quickly attracted to the intracellular space. This is called excitatory post-synaptic potential (EPSP). When there are sufficient number of EPSPs that occur in close succession at the same region of the synaptic membrane (to allow temporal and spatial summation of EPSPs; see **Figure 4-5**), the depolarization becomes strong enough to travel down to the axon hillock and trigger the action potential (see below).

Hyperpolarization can counteract depolarization. Hyperpolarization occurs when ion channels that allow influx of negatively charged ions (or in some cases, efflux of positively charged ions down their concentration gradient) are opened. The opening of these channels is triggered when certain neurotransmitter substances bind to specific receptors. Hyperpolarization is also called inhibitory post-synaptic potential (IPSP). IPSPs often target the cell body rather than the distal dendrites, thus more effectively preventing EPSPs induced by depolarization from moving down from distal dendrites to the axon hillock.

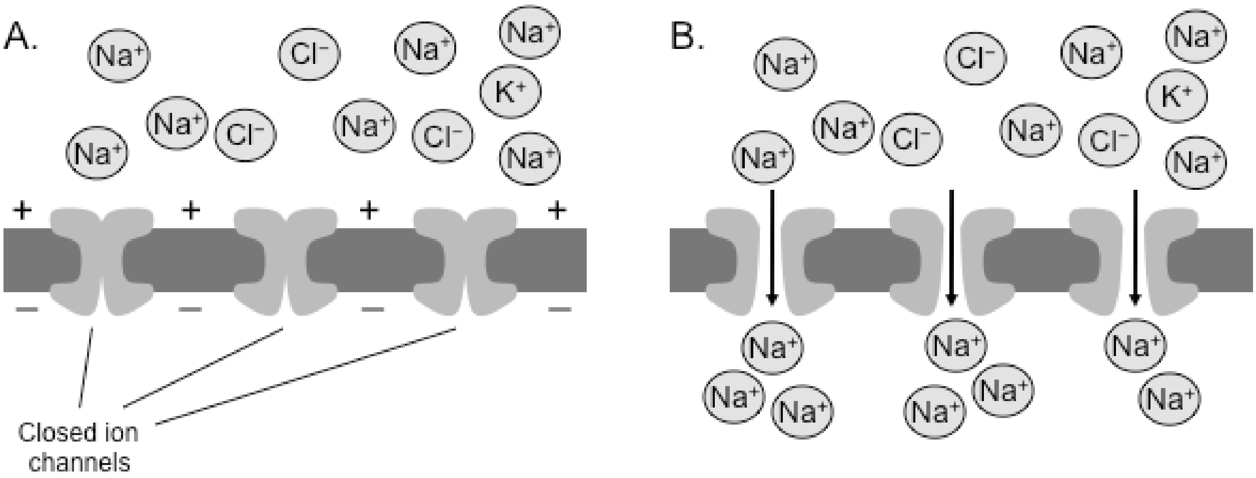


Figure 4-4. Schematic illustration of what happens during depolarization. A) In the resting state, channels that would allow influx of positively charged ions (that are attracted to intracellular space due to the voltage difference) are closed. B) Opening of these ion channels leads to influx of positively charged ions and depolarization of the neuronal membrane (*i.e.*, the negative difference between intra- and extracellular space is reduced). When sufficient number of such depolarizing excitatory post-synaptic potentials accumulates temporally and spatially, the depolarization spreads down to axon hillock where the action potential is initiated.

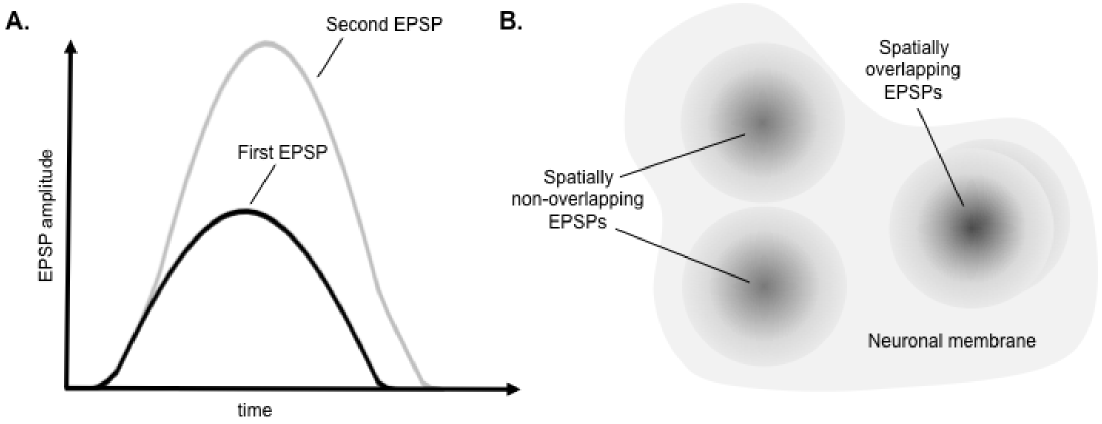


Figure 4-5: Temporal and spatial summation of EPSPs makes it possible for the depolarization to become strong enough to travel down from the dendrites to axon hillock where the action potential is triggered. A. Temporal summation of two EPSPs that overlap spatially is shown. B. Spatial summation of temporally coincident EPSPs is shown.

Generation of the all-or-none action potential

When the membrane potential difference is reduced to –50 mV (in typical neurons that have resting membrane potential of –70 mV) in the axon hillock due to accumulation of depolarization caused by EPSPs, the so-called voltage-gated ion channels are transiently opened. This results in quick influx of Na^+ and/or Ca^{2+} and the membrane potential quickly becomes positive (as opposed to the negative resting membrane potential). This phenomenon is highly transient as the voltage-gated channels are quickly closed and the resting membrane potential is restored by the Na^+/K^+ pump.

Once generated the action potential travels down the length of the axon to terminal endings in an all-or-none fashion. In the terminal endings release of neurotransmitter to dendrites of the post-synaptic cells is then triggered. An action potential is followed by a brief refractory period during which another action potential cannot be generated, thus limiting the firing rate of a neuron to about 1 kHz. The central principles of the chemical messaging in the synapses are described next.

4.3 Chemical messaging at the synapses

There are a number of steps in chemical messaging over the synaptic cleft. While most often the terminal endings release transmitter onto the dendrites of the post-synaptic cell, there are also synapses onto the cell body and even axons of other neurons. An idealized synapse presenting the essential steps of chemical messaging is presented in **Figure 4-6** below. Chemical messaging at the synapses constitute targets of most of the drugs that are designed to affect central nervous system function (Cooper et al., 2003).

Neurotransmitter is released into the synaptic cleft

Chemical messaging begins when the action potential travels down the axon to the terminal endings. Neurotransmitter that is synthesized by the presynaptic neuron and stored in vesicles is then released to the synaptic cleft, which is a small gap between the pre- and post-synaptic neurons. After being released, the transmitter molecules diffuse across the synaptic cleft and bind to specific receptor molecules of the post-synaptic cell. Notably, while a single presynaptic neuron typically does not release more than one type of transmitter substance, there often are multiple types of receptors on the cell membrane of a post-synaptic neuron.

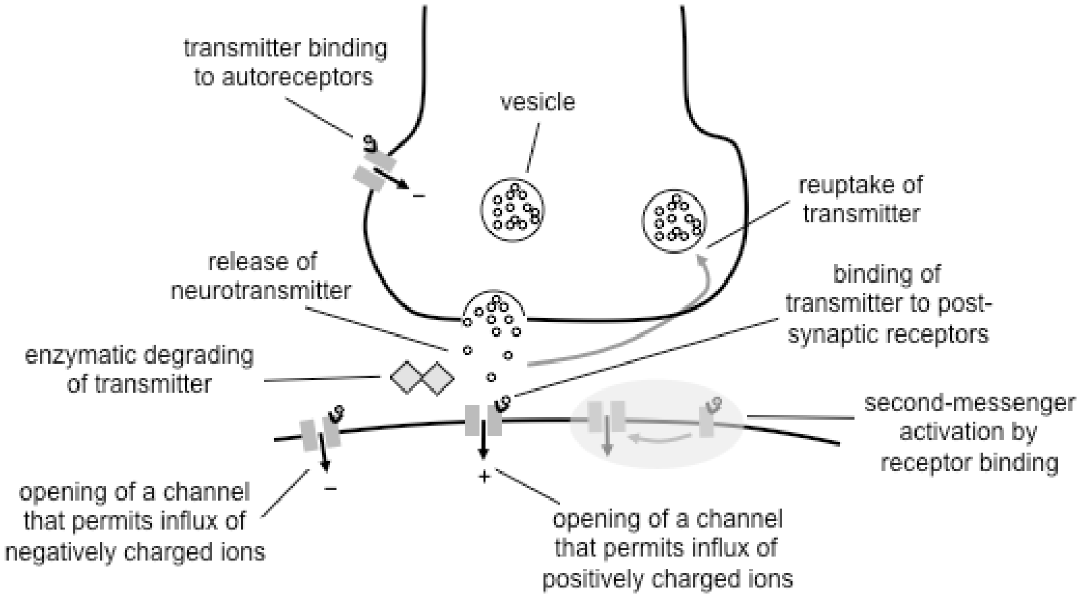


Figure 4-6. Shown is a schematic illustration of an idealized synapse. The terminal ending of the pre-synaptic neuron (depicted on top) contains vesicles that store and release neurotransmitter to the synaptic cleft. Neurotransmitter then binds to the receptors that are located on the membrane of (typically dendrites) of the post-synaptic neuron. The different steps of chemical transmission at the synapse are shown in this figure. See text for details.

Neurotransmitters exert their effects by binding to receptors

The type of the receptor dictates what kind of effect the neurotransmitter has on the post-synaptic neuron. There are receptors that are directly coupled to ion channels. Such receptors can induce very quickly either depolarization (when coupled to ion channels that allow influx of positively charged ions) or hyperpolarization (when coupled to ion channels that allow influx of positively charged ions). In addition to these receptors, there are receptors that modulate ion channels indirectly *via* so-called second messenger mechanisms. The second-messenger mediated effects are slower and longer lasting than those caused by receptors that are directly coupled to ion-channels. Autoreceptors constitute yet another type of receptor that is shown in **Figure 4-6** above. The role of the autoreceptors in general is to regulate/inhibit further release of neurotransmitter by the presynaptic cell.

Central nervous system drugs often target specific receptors

Quite often drugs affecting the central nervous system target specific receptors. By binding to a specific receptor, drugs can induce the same effect as endogenous neurotransmitter molecules, enhance the effects of neurotransmitters, or block the receptors and thus prevent neurotransmitters from binding to the receptor and the ion channels from opening. Drugs acting alike neurotransmitters are called agonists and drugs that block receptors are called antagonists. A further distinction is made between competitive and non-competetive antagonists, which reflects the potency of the drugs to replace transmitter from the receptor (*i.e.*, competitive antagonists replace endogenous transmitter from receptors). Note also that antagonism of autoreceptors leads to increase in the amount of transmitter in the synaptic cleft by reducing the feedback-inhibition effects on transmitter release.

Neurotransmitters are actively removed from the synapse

Neuroransmitter that has been released is cleaned away from the synapse by two mechanisms: there is active re-uptake of transmitter molecules back to the presynaptic neuron and there are specific enzymes that degrade the transmitter (see **Figure 4-6**). Modern antidepressant drugs act by inhibiting the re-uptake of a specific neurotransmitter called serotonin (also known as 5-hydroxytryptamine) back to the presynaptic cell, thus alleviating serotonin deficiency that clinical depression has been associated with. Acetylcholinesterase degrades a neurotransmitter called acetylcholine. Given that acetylcholine is lacking in patients suffering from Alzheimer's disease (one of the major dementias), acetylcholinesterase inhibitors have proved effective in improving the condition of these patients by increasing the amount of acetylcholine in the synapses (Birks, 2009).

The major neurotransmitters in the brain

The most common neurotransmitters in the brain are amino acids glutamate and gamma-aminobutyric acid (GABA). Glutamate is the most common excitatory neurotransmitter that binds to multiple receptor types, perhaps the most well known of them being the N-methyl-D-aspartate (NMDA) receptors (for a recent review on the subtypes of NMDA receptors, see (Paoletti, 2011)). As will be reviewed in Chapter 8, the NMDA receptors are highly important for learning and plasticity. GABA, in turn, is the most common inhibitory neurotransmitter. There are different types of GABA receptors. The most well known are the $GABA_A$ receptor that is directly coupled to an ion channel that, when opened, allow influx of Cl^- to the cell, thus leading to hyperpolarization, and the $GABA_B$ receptors that operate with a second messenger mechanism (for a review, see (Bormann, 1988)). In the developing nervous system, as well as in specific areas of the adult brain such as striatum, however, GABA causes depolarization (for a review, see (Ben-Ari, 2002)).

Acetylcholine, dopamine, norepinephrine, and serotonin are highly important and widely investigated neurotransmitters that play key roles in cognitive functions such as learning, memory, and selective attention. Acetylcholine was the first neurotransmitter that was found. In early studies, it was observed that release of acetylcholine by motor neurons is what causes muscles to contract. In the brain, acetylcholine is highly important for plasticity and formation of new memories (Froemke et al., 2007). The catecholamines dopamine and norepinephrine are, in turn, important for attentional functions.

Serotonin (also called 5-hydroxytryptamine or 5-HT) has been linked to regulation of arousal and its deficiency plays a central role in mood disorders (for a review, see (Nemeroff and Owens, 2002)). In addition, there are numerous other neurotransmitter substances including peptides that have a central influence on cognition. As an example of this, endogenous opiates (also called opioids), such as beta-endorphin, have a profound influence on emotions, motivation, and have been suggested to play a key role in the development of addictions such as alcoholism (for a review, see (Sinclair, 2001)).

4.4 Neurons form networks

The principles outlined above of how neurons can modulate the activity of other neurons form the basis for functional neuronal networks. There are a few principles and phenomena that are relevant for understanding how networks of neurons function. For example, in a case where only a few sensory neurons are firing when a spot of light is shined on the retina, the output of these neurons need to be amplified in order for them to influence sufficient number of neurons in the brain. On the other hand, under poorly illuminated conditions, the input from multiple sensory neurons needs to be integrated in order to detect faint visual stimuli. These two phenomena, referred to as convergence and divergence, are described below. Other network-level phenomena that are relevant for perceptual and cognitive functions include oscillatory activity of neuronal populations, and how neuronal populations modulate each other *via* inhibitory and excitatory connections to regulate behavior, a good example of which is offered by the interconnectivity of basal ganglia in regulating selection of movements and cognitive sets (see Chapter 10).

Convergence and divergence

Convergence and divergence simply refer to integration of inputs from multiple neurons and distribution of a signal to large population of neurons, respectively. This is illustrated in **Figure 4-7** below. Good examples of convergence can be found in the retina of the human eye, where so-called rod cells that underlie peripheral vision converge their inputs onto a ganglion cell. This makes it possible for the rods to be very sensitive to light and enable one to see in dark. However, the downside is that visual acuity is very low when there is very little light (you can try this yourself by trying to read a book in twilight: while you can see the book and the text it is not possible to discern the individual letters).

Divergence, on the other hand makes it possible for a significant signal emitted by small amount of neurons to be amplified in the brain. Example of this is how light stimulating a single cone cell in the fovea of retina (fovea refers to the small area in the retina that makes it possible to see accurately and perceive colors) gets processed by multiple brain areas in the visual system. Large-scale brain connectivity presents another example of divergence, as primary sensory areas signal multiple secondary sensory areas and thus spread the signal.

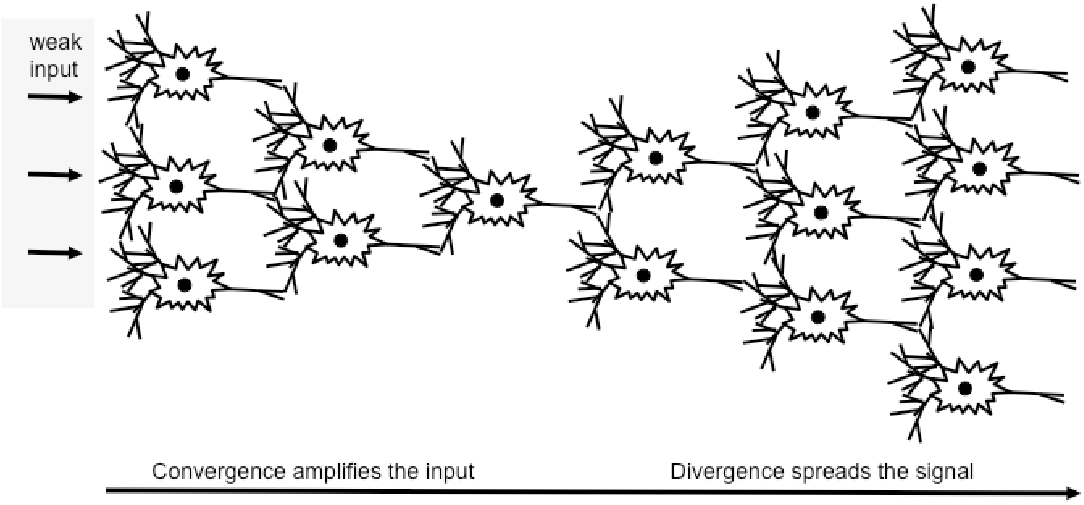


Figure 4-7: Convergence and divergence of neural inputs/signals helps amplify and spread the signal, respectively. See text for details.

Inhibition and excitation in the regulation of behavior

The fact that neurons can exert inhibitory and excitatory effects on other neurons plays a central role in the organization of functional networks of neurons. The stretch reflex provides a simple example of how excitation and inhibition can work in concert as a simple neuronal network (see **Figure 4-8**). When the knee is hit with a hammer, the stretch causes specific neurons to depolarize, sending action potentials to spinal cord. Within the spinal cord, these inputs depolarize motor neurons that innervate the extensor muscles of the leg, thus causing the leg to contract.

However, if the antagonist flexor muscle would not be relaxed at the same time, effort would be spent uselessly. This is solved by the stretch-sensitive neuron exciting an inhibitory interneuron within the spinal cord that then inhibits the motor neurons that innervate the flexor muscle. Naturally, the interplay between excitation and inhibition is a lot more complex in the brain than in the spinal reflex. One good example of this is the connectivity of the basal ganglia, a set of nuclei that regulate initiation of motor responses and changes in so-called cognitive sets (for details, see Chapter 10).

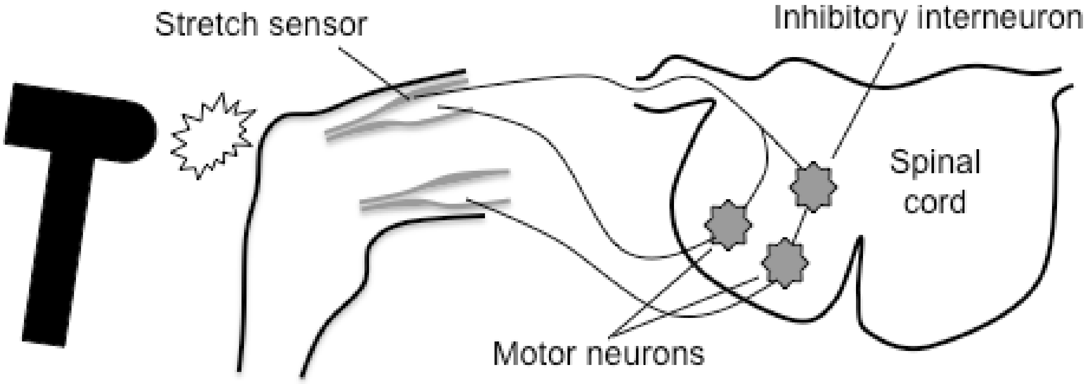


Figure 4-8. The stretch reflex that results from a light tap of hammer under the knee involves a simple neuronal network of inhibition and excitation. See text for details.

Population-level oscillatory phenomena

Oscillation of neural populations at various frequencies, which seems to be enabled by inhibitory interneurons (Chen and Dzakpasu, 2010), is a phenomenon that has elicited a lot of interest in cognitive neuroscience due to hypothesized links between consciousness and oscillatory activity (see Chapter 13). Much like divergence of signals described above, temporal synchrony (*i.e.*, phase locking) of activity of multiple neurons allows emergence/augmentation of a population signal (due to temporal synchrony).

Indeed, it has been suggested that phase locking of oscillatory activity between brain areas could underlie the so-called perceptual binding problem (Bertrand and Tallon-Baudry, 2000). This refers to the problem of how parallel processing of various features of perceptual objects (such as the red color, characteristic shape, and texture of an apple) that takes place in anatomically distinct areas is bound together to allow coherent percept of the object. Phase locking of oscillatory activity across distinct brain areas where the relevant processing takes place has been proposed as the candidate underlying neural mechanism.

4.5 Concluding remarks

Neurons are the elementary building blocks of cognition. Ultimately, it is the vast number of neurons contained in the 1.4 kilograms of tissue packed inside the human skull that, through dynamic interactions and networking, make perception, cognition, emotion, and consciousness possible. There are a few fundamental mechanisms that underlie the functioning of neurons. Maintaining the resting membrane potential makes it possible for depolarization, and thus neuronal signaling, to occur quickly. On the other hand, entire populations of neurons can be inhibited from firing action potentials with release of neurotransmitters such as GABA that result in hyperpolarization.

Indeed, chemical messaging at the microscopically small contact points between pre- and postsynaptic neurons, synapses, are specific targets for a plethora of central nervous system drugs that help ameliorate symptoms of devastating disorders such as dementias of Alzheimer and Parkinson's type, debilitating mental health disorders such as schizophrenia and depression, as well as developmental disorders such as attention deficit hyperactivity disorder. In the next chapter of this book, it is described how neurons operating within the sensory and motor systems make it possible for humans to see, hear, feel, and move about.

4.6 References

BEN-ARI, Y. 2002. Excitatory actions of GABA during development: the nature of the nurture. *Nature Reviews Neuroscience,* 3, 728-739.

BERTRAND, O. & TALLON-BAUDRY, C. 2000. Oscillatory gamma activity in humans: a possible role for object representation. *International Journal of Psychophysiology,* 38, 211-223.

BIRKS, J. 2009. *Cholinesterase inhibitors for Alzheimer's disease*, The Cohcrane Library.

BORMANN, J. 1988. Electrophysiology of GABA-A and GABA-B receptor subtypes. *Trends in Neurosciences,* 11, 112-116.

CHEN, X. & DZAKPASU, R. 2010. Observed network dynamics from altering the balance between excitatory and inhibitory neurons in cultured networks. *Physical Review E,* 82, 031907.

COOPER, J. R., BLOOM, F. E. & ROTH, R. H. 2003. *The biochemical basis of neuropharmacology,* New York, Oxford University Press.

FROEMKE, R. C., MERZENICH, M. M. & SCHREINER, C. E. 2007. A synaptic memory trace for cortical receptive field plasticity. *Nature,* 450, 425-429.

HANISCH, U.-K. & KETTENMANN, H. 2007. Microglia: active sensor and versatile effector cells in the normal and pathologic brain. *Nature Neuroscience,* 10, 1387-1394.

HARTLINE, D. K. & COLMAN, D. R. 2007. Rapid conduction and the evolution of giant axons and myelinated fibers. *Current Biology,* 17, R29-R35.

JANZER, R. C. & RAFF, M. C. 1987. Astrocytes induce blood-brain barrier properties in endothelial cells. *Nature,* 325, 253-257.

NEMEROFF, C. B. & OWENS, M. J. 2002. Treatment of mood disorders. *Nature Neuroscience,* 5, 1068-1070.

PAOLETTI, P. 2011. Molecular basis of NMDA receptor functional diversity. *European Journal of Neuroscience,* 33, 1351-1365.

SINCLAIR, J. D. 2001. Evidence about the use of naltrexone and for different ways of using it in the treatment of alcoholism. *Alcohol Alcohol,* 36, 2-10.

5 Sensory and motor systems

What is striking about the processing of sensory stimuli (*i.e.*, sights, sounds, smells, and tactile stimuli) and executing / coordination of motor actions (*i.e.*, movements of one's body when one stands, walks, reaches etc) is that the majority of brain tissue seems to be participating in these tasks. Without knowing this, especially if one is healthy and do not have perceptual/motor problems, one might think that perceiving the world and moving about is a fairly trivial task. However, the enormous complexity of these functions is exemplified by the fact that it has been extremely difficult for engineers to design robots that could perceive the world, walk about, and reach for/use objects as humans do. What makes the sensory and motor systems even more interesting for a cognitive neuroscientist is that the very brain areas that are involved in processing sensory stimuli and carrying out motor actions are also utilized when one imagines, thinks, and plans ahead what one will do next. In the following, the sensory systems are described first followed by an introduction to the motor system.

5.1 The sensory systems

The sensory systems consist of visual, auditory, somatosensory, and olfactory systems. Notably, while this division is based on the sense organs, there are multiple receptor types that sense different aspects of sensory stimuli within each of the senses. As examples of this, there are different receptors in our skin for sensing touch/texture and pain. There are four different types of receptors in the eye that allow perception of different aspects of visual stimuli such as color, fine detail, and movement. In the following, the neural mechanisms of vision are described first, followed by brief introductions to the auditory, somatosensory, and olfactory pathways of the brain.

Receptors in the retina convert light to neural impulses

The most essential parts of the eye and the retina that contains the receptor neurons that are sensitive to light are illustrated in **Figure 5-1**. Light that hits the eye is focused and projected to the back wall of the eyeball through the lens. The retina covers the back of the eye. It is translucent and the light-receptor cells, called rods and cones, are located at the very back of the retina so that light travels through the retina before hitting the receptors. There are altogether four types of light-sensitive cells in the human retina: three types of cone and a single type of rod cells.

The rods are larger and more sensitive to light than the cones and occupy the parts of retina that receive light from the periphery of the visual field. The cones are concentrated at the so-called fovea (*i.e.*, a point on the retina where there is high visual acuity). The cones and rods send their input to ganglion cells through a layer of cells containing horizontal, bipolar, and amacrine cells. Whereas rods are all sensitive to same spectrum of visible light wavelengths, the three subtypes of cones are sensitive to distinct ranges of wavelengths.

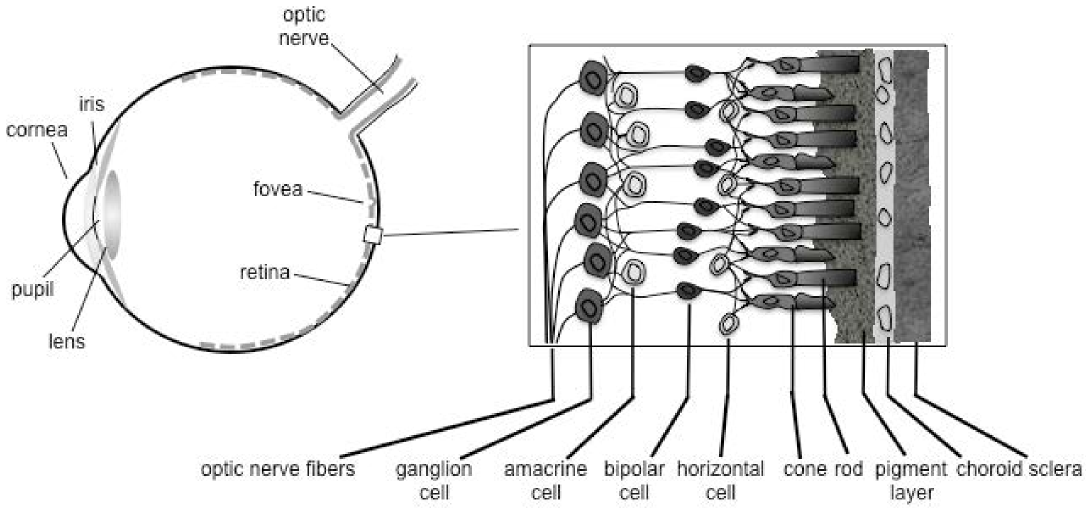


Figure 5-1: The human eye (left) and cell structure of the retina enlarged (right). See text for details.

This differential sensitivity of the cones to different wavelengths (see **Figure 5-2**) underlies the ability to perceive colors. On the other hand, the enhanced light sensitivity of the rods allows one to see in darkness, although the rods do not allow as high visual acuity as the cones due to convergence of their inputs to ganglion cells that form the optic nerve (see Chapter 4 for the description of convergence). Cones on the other hand can directly send input to a single ganglion cell, thus allowing for a very high visual acuity.

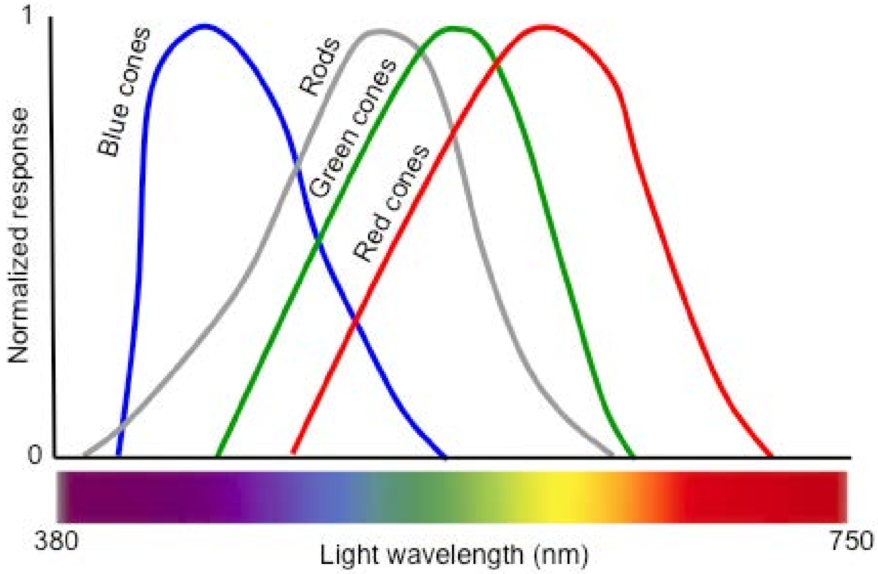


Figure 5-2. Shown is the differential sensitivity of rods and the three types of cones to visible light spectrum. Note that normalized sensitivity as a function of wavelength of light is plotted here. The rods are the most, and the blue cones the least, sensitive to light.

The visual pathway: from retina to primary visual cortex

The axons of the retinal ganglion cells form the optic nerve that exits from the eye and conveys visual information first to lateral geniculate nucleus (LGN) of the thalamus. While there are some afferents to superior colliculi that enable rudimentary visual functions (*e.g.*, localization of stimuli) without conscious awareness ("blindsight") in patients with blindness due to visual cortex lesions (Weiskrantz, 1996), most of the fibers go to LGN. From LGN visual information is passed on to the primary visual cortex that resides within the calcarine fissure. Notably, visual inputs cross from the left hemifield to the right LGN and right visual cortex, and *vice versa* (see **Figure 5-3**). This crossing of inputs is common to all the sensory pathways with the exception of the olfactory system.

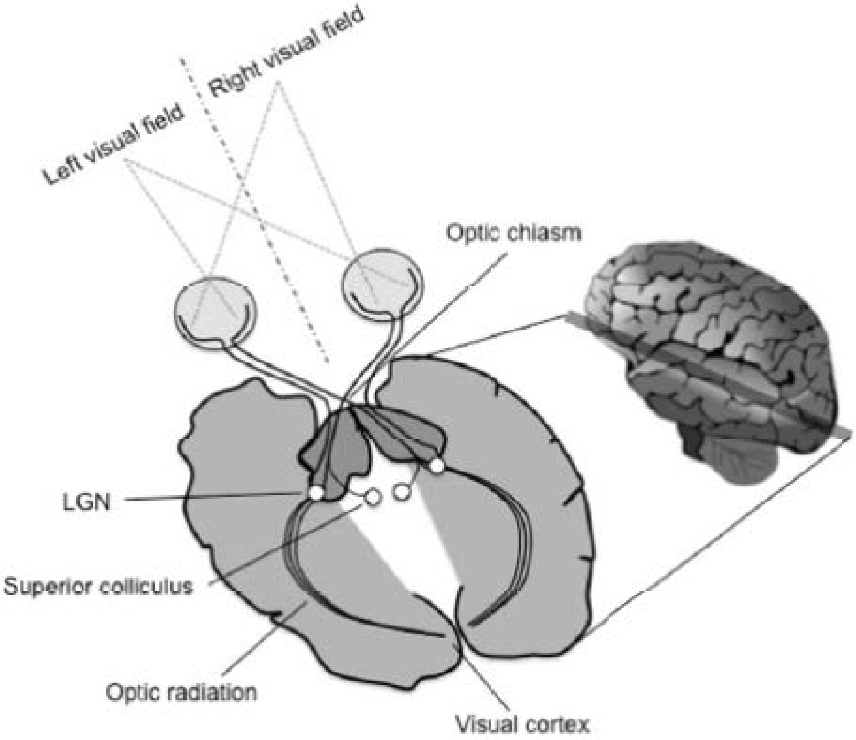


Figure 5-3. Visual information that is converted by the rods and cones in the retina to neural impulses travels to the visual cortex in the occipital lobe *via* the lateral geniculate nucleus (LGN) of the thalamus. There is a smaller subcortical connection to superior colliculus that has been associated with rudimentary visual abilities without conscious awareness of visual stimuli in patients who have become blind due to lesions of the visual cortex.

Magno- and parvo- and koniocellular systems

Within the LGN, visual inputs from the ganlion cells terminate into six layers. Three of these layers are innervated by the inputs from the left, and three from the right, eye. One of these three layers is targeted by inputs from ganglion cells that are stimulated by rods and two of the three layers are targeted by ganglion cells that are innervated by cones. The two layers targeted by rod-innervated ganglion cells are called magnocellular due to the larger size of the neurons and the four layers innervated by cones are called parvocellular.

In addition to the magno and parvocellular layers, thinner layers between the magno- and parvocellular layers have been described called koniocelllular layers (Stewart et al., 2000). These cells receive their inputs from the blue-sensitive cones.

Most of the inputs that the LGN receives are, however, feedback connections from the visual cortex and connections from the reticular activation system. These connections may help the visual system to filter task-relevant stimuli for further processing (see Chapter 6).

Visual cortex neurons exhibit a hierarchy of receptive field complexity

The now-classic observations of Hubel and Wiesel (Hubel and Wiesel, 1959) of the type and hierarchy of receptive fields of visual cortex neurons were awarded with the Nobel Prize in 1981. What Hubel and Wiesel observed was that some neurons in the visual cortex responded to lines of a specific orientation. They named these cells the "simple cells", to dissociate them from hierarchically higher-order ("complex" and "hypercomplex") cells that respond to contrast edges moving to specific directions (Hubel and Wiesel, 1968). Hubel and Wiesel also observed that the primary visual cortex is organized in columns. Every other column is innervated by inputs from left and every other by inputs from the right eye. Each of these ocular dominance columns, in turn, contains orientation columns that form an orderly gradient of orientation selectivity (*i.e.*, from horizontal through tilted and vertical back to horizontal).

Perceptually, it is thought that the receptive fields of the simple cells are reflected in that we detect very well contrast edges. For instance, a blank wall does not contain much information that would capture our attention whereas a doorframe is very quick to detect due to the visual system being sensitive to contrast edges between the doorframe and the wall. Higher up the in visual system the receptive fields of neurons do not become excessively complex (*i.e.*, there are no "grandmother" cells that would only be selective to perceiving one specific object), but rather it seems that perceptual objects are coded by distributed firing patterns of populations of neurons (Ishai et al., 1999). For an example of such findings, see Figure **5-4** below.

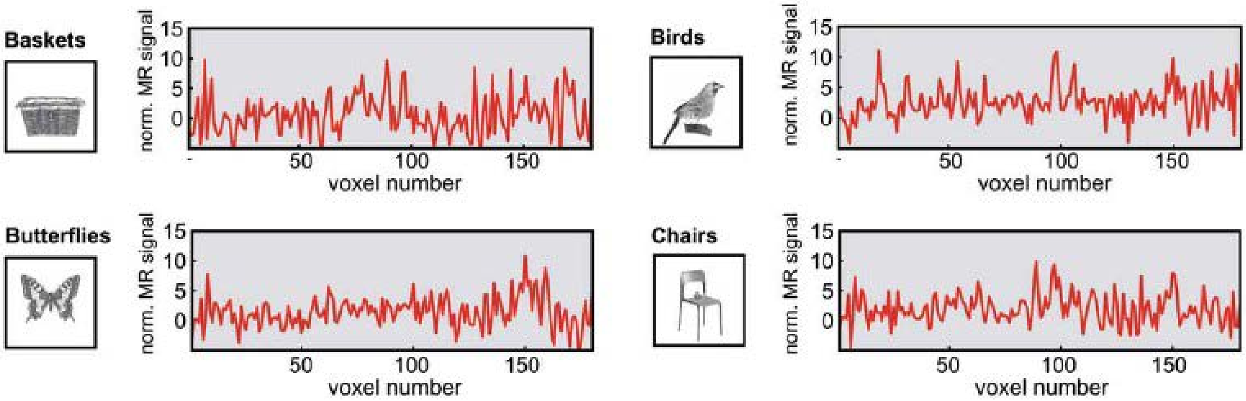


Figure 5-4: Shown are distributed representations of perceptual objects that were observed, using specific machine-learning algorithms, in fMRI data obtained in healthy volunteers during viewing of the objects. Here are shown "signature patterns" of fMRI response magnitude across a set of voxels covering the lateral occipital complex, observed during viewing of baskets, birds, butterflies, and chairs (adapted with permission from (Cox and Savoy, 2003)).

Visual cortical areas are retinotopically organized

Functionally distinct visual cortical areas have been identified by mapping separate retinotopically organized areas in the visual system. The concept of retinotopic map refers to the fact that the visual inputs from each hemifield are projected to adjacent loci on the cortical surface of the opposite hemisphere. In such maps, the representation of the fovea is disproportional: the fovea takes up only one percent of retinal size, but up to 50% of the visual cortex processes foveal inputs. There are multiple retinotopically organized areas in the human occipital lobes (DeYoe et al., 1996), with each devoted to processing of different aspects of visual input.

The "what" and "where/how" processing pathways in vision

Features that neurons in each of the visual areas are sensitive to vary considerably. While ventral visual areas have been implicated in processing of object identity (color, texture, shape, and identity), there is a continuum of areas when progressing dorsally from the primary visual cortex that seem to be devoted to processing of movement and spatial location of the visual objects. This division of processing to ventral and dorsal streams has been coined the "*what*" and "*where*" streams, based on the presumed functional roles of the pathways (Mishkin and Ungerleider, 1982). Recent evidence suggests that the dorsal stream is also involved in spatial-motor processing, for instance, perception of the actions of other people (Goodale and Milner, 1992), and thus it has also been called the dorsal "*how*" stream. The dorsal and ventral streams are schematically illustrated on the surface of the brain in **Figure 5-5** below.

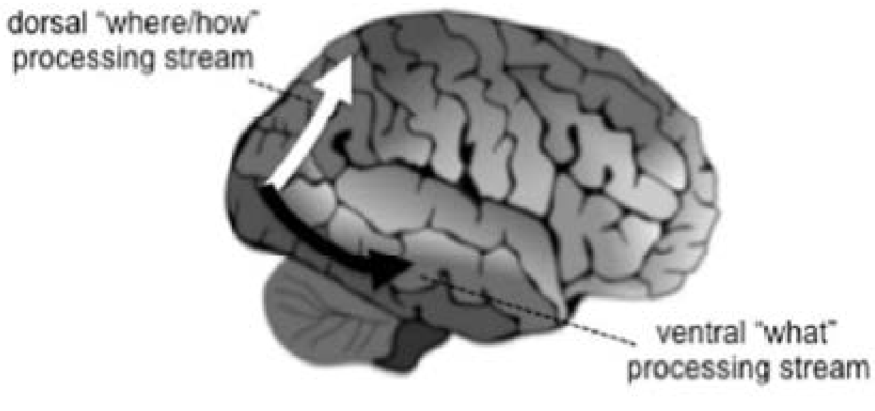


Figure 5-5: The dorsal "where/how" and ventral "what" visual processing streams (shown with the white and black arrows, respectively, plotted on the right hemisphere surface) that process the spatial locations and identities of visual objects, respectively.

In addition to evidence obtained in neuroimaging studies where differential activity patterns have been observed during object identity *vs.* spatial location tasks and manipulations, the segregation of object identity and stimulus movement/spatial location processing is also manifested in neurological patients with focal brain lesions. Akinetopsia refers to a rare condition where the patient is unable to detect movements of stimuli and objects (for a review, see (Zeki, 1991)). Agnosia on the other hand refers to conditions where the patient is unable to recognize objects. For example, a patient, while still being able to draw a phone or a dog, might be unable to name what there is in the picture when it is later shown to him/her (for a review of the different types of agnosias, see (Biran and Coslett, 2003)).

Hair cells of basilar membrane transform sounds into neural impulses

Whereas retinal rods and cones transform light into neural impulses, there are specific hair cells in the inner ear that transform sounds into neural impulses (*i.e.*, physical bending of the hair cells causes them to depolarize and fire). These hair cells are located on a membrane called basilar membrane that is coiled inside a structure called cochlea (see **Figure 5-6**). Sounds arriving to the ear first resonate the so-called tympanic membrane. This resonance is then transferred *via* three small bones called malleus, incus, and stapes, to the cochlea and onto the basilar membrane. Within the cochlea the basilar membrane vibrates frequency-specifically; low sound frequencies resonate basilar membrane at the apex, mid-frequency sounds in the middle, and high frequency sounds at the base of the cochlea. This is due to the base of the basilar membrane being broader than its tip.

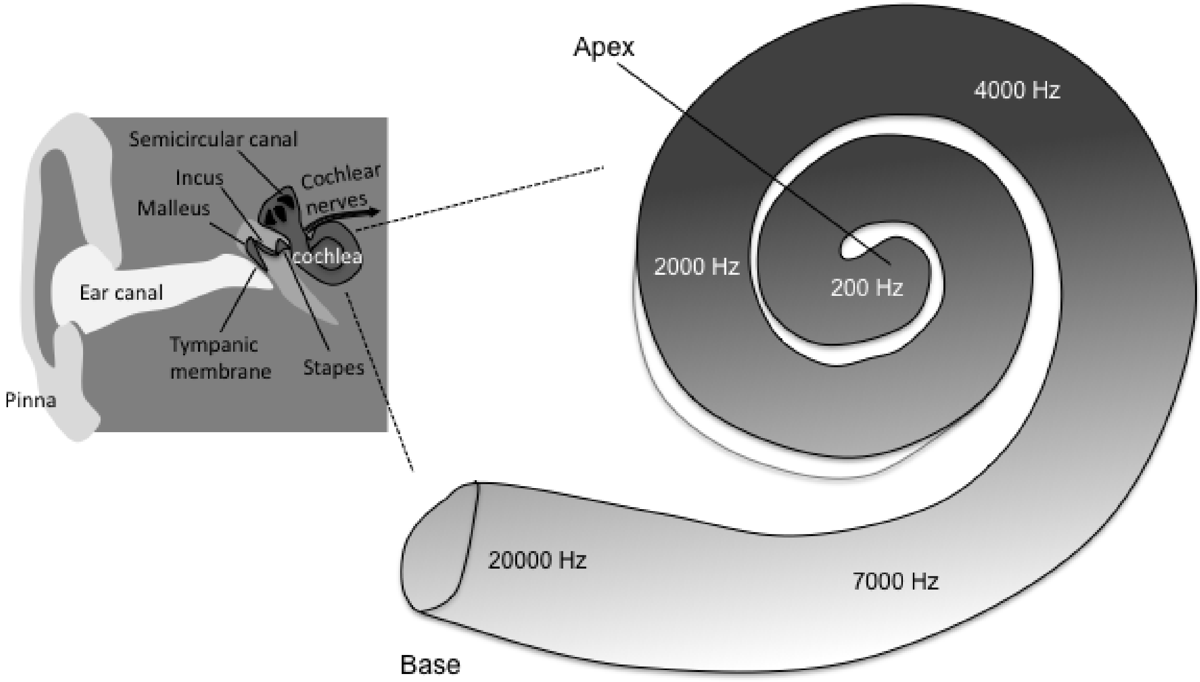
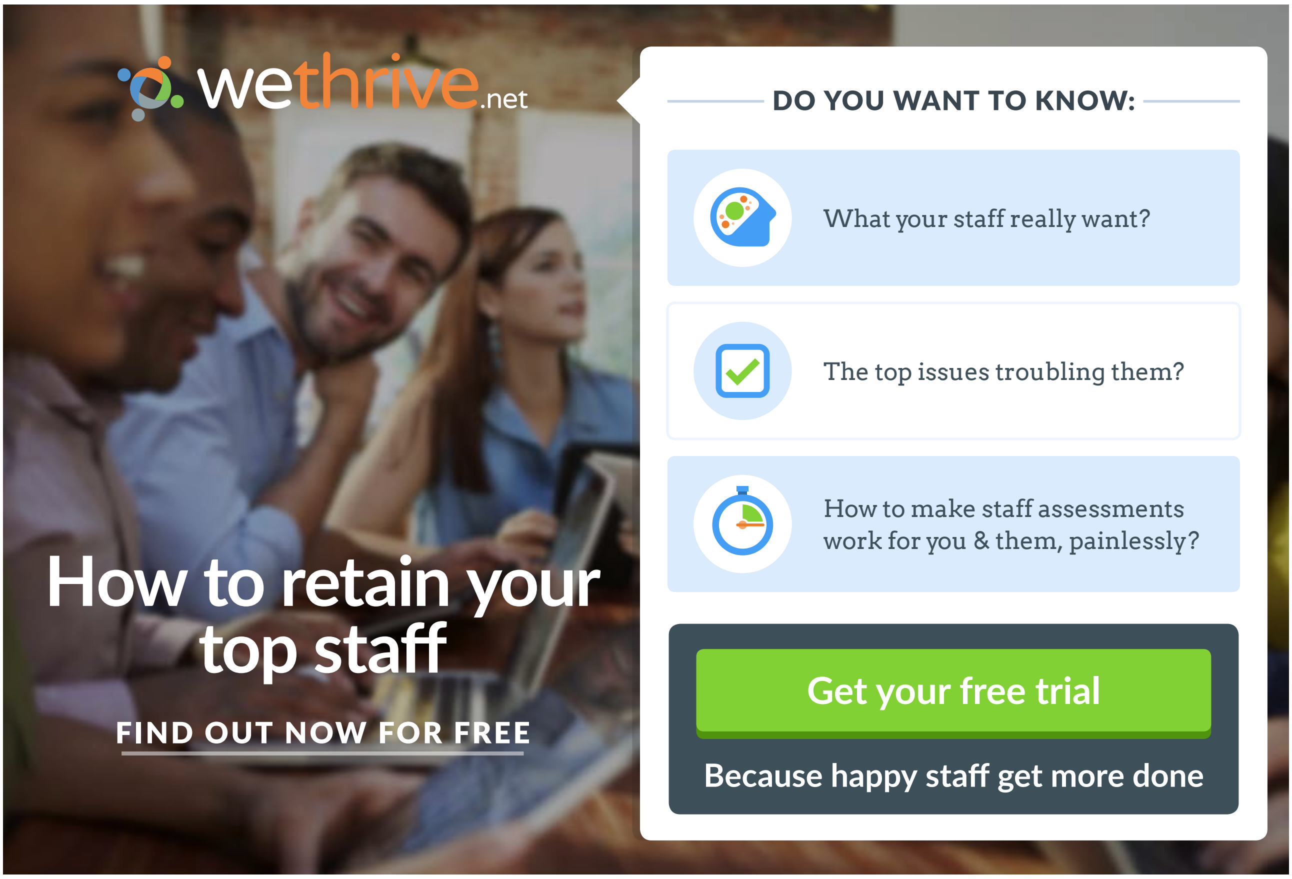


Figure 5-6. Shown are the structures of the human ear in a schematic illustration, with the cochlea and the gradient of sound frequency sensitivity of the basilar membrane enlarged. At the base of the cohclea the basilar membrane is selective to high sound frequencies (up to 20 kHz) and at the apex to low sound frequencies (200 Hz).



The differential resonance characteristics of the basilar membrane creates a place code that represent sound frequencies in auditory cortical areas that reside on the superior aspect of the temporal lobe bilaterally. There are multiple tonotopic representations, which means that there is an orderly progression, along the cortical surface, of neurons responding from low to high sound frequencies (Talavage et al., 2004). While the term tonotopic organization is most widely used in cognitive neuroscience, it can be argued whether "cochleotopic organization" could be an even more fitting term (*i.e.*, different parts of the basilar membrane being spatially represented on the cortical surface similarly to retinotopic organization in the visual system).

The place code effectively allows the auditory system to decode the spectrotemporal properties of sounds, thus enabling efficient identification of highly complex natural sounds (Lewicki, 2002). In addition to the place code, the firing frequency of neurons within the auditory nerve is though to help represent sounds. Even though the firing rate of a single neuron can only approach 1 kHz, it has been proposed that neurons fire in volleys thus constituting a frequency code that further helps represent sounds.

Input from haircells travels through multiple subcortical nuclei to cortex

Much like in the case of the visual system, there are subcortical nuclei through which auditory inputs traverse to the primary auditory cortex (see **Figure 5-7**): cochlear nucleus, superior olive, inferior colliculus and the medial geniculate body of the thalamus. The primary auditory cortex is located within the Sylvian fissure on the medial aspect of the Heslch's transverse gyrus at the superior temporal lobe. The receptive fields of the neurons become more complex as one ascends the auditory pathway. For instance, in animal studies, selective responses to features of species-specific vocalizations have been documented in the inferior colliculus and some researchers have equated the complexity of auditory feature processing of inferior colliculus with that of the primary visual cortex (Nelken et al., 2003).

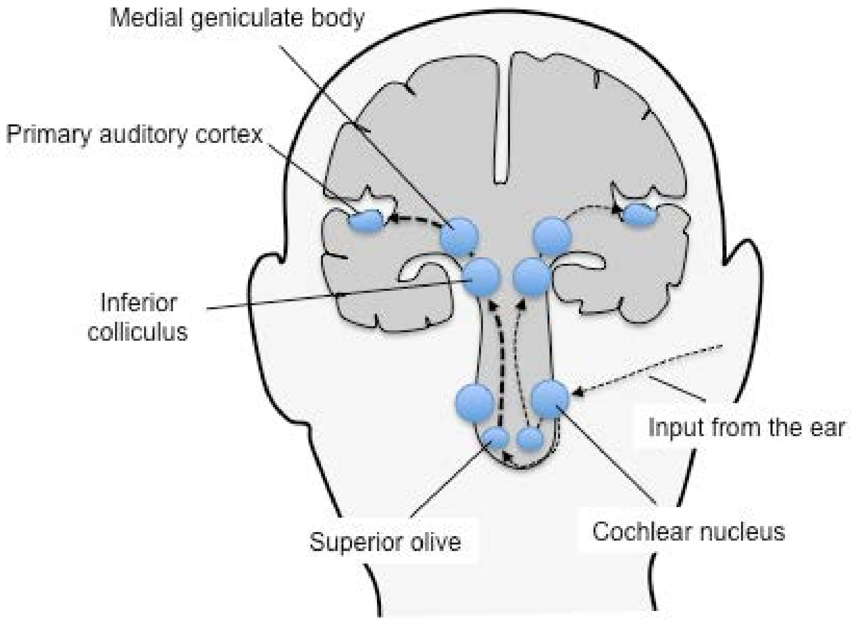


Figure 5-7: The auditory pathway. Sounds are converted in the cochlea to neural impulses, and the neural impulses travel through several nuclei (cochlear nucleus, superior olive, inferior colliculus and medial geniculate body of the thalamus) to primary auditory cortex that occupies the medial two-thirds of the Heschl's gyrus in the superior aspect of the temporal lobe. Note that, while most of the fibers cross from the lear ear to the right auditory cortex and vice versa, there are also ipsilateral fibers that ascend from the ear to the auditory cortex on the same side.

In addition to the connections of the auditory pathway shown in **Figure 5-7**, there are connections from inferior colliculus to amygdala that play a central role in emotional processing (see Chapter 11). Furthermore, descending connections (that are mainly ipsilateral for some reason) are as much as an order of magnitude more abundant than the ascending connections. This has been suggested to be a potential mechanism that facilitates attending selectively to certain sounds at the expense of task-irrelevant sounds by way of enhancing the processing of attended sounds (see Chapter 6).

Hierarchical processing within auditory "what" and "where/how" streams

There are several secondary auditory cortical areas that surround the relatively small primary auditory cortex that occupies the medial two-thirds of the Helschl's gyrus on the superior aspect of the temporal lobe. These areas are hierarchically organized: whereas more simple auditory stimulus features/stimuli such as tones elicit responses mainly in the primary auditory cortex, secondary auditory cortex areas respond to increasingly complex sounds such as broadband noise bursts, speech sounds, and words (Binder et al., 2000). Further, similarly to the visual system there is evidence for parallel "*what*" and "*where*" streams, with anterior/ventral secondary auditory cortex areas responding to speech sounds (or in monkeys species specific vocalizations) and posterior/dorsal areas selectively responding to spatial locations of sounds (Rauschecker and Tian, 2000, Ahveninen et al., 2006).

The dorsal "*where*" processing pathway traverses from the auditory cortex through temporal-parietal junction to posterior parietal cortex and ultimately to dorsal prefrontal cortical areas. The ventral "*what*" pathway progresses to ventral-anterior direction from the primary auditory cortex and reaches the inferior lateral prefrontal cortical areas through a connection called the uncinate fasciculus. While the dorsal stream was originally implicated to be involved in processing of spatial locations, recent studies have suggested that the dorsal stream is also involved in processing of "*how*" information (*e.g.*, mapping speech auditory inputs to speech motor sequences) (Hickok and Poeppel, 2004, Rauschecker and Scott, 2009).

Somatosensation: senses of touch, pain, and temperature

The senses of touch, pain (nociception), body posture (proprioception), and the ability to feel temperature are all based on specific receptors in the skin. Together this collection of senses has been termed somatosensation and the system as a whole the somatosensory system. The somatosensory receptors that are located in the skin, joints, and various internal organs are called mechanoreceptors, nociceptors, and thermoreceptors. Mechanoreceptors are a class of receptors that convert mechanical pressure or distortion of skin to neural impulses. There are four types of mechanoreceptors in the human skin: Pacinian, Meissner's and Ruffini corpuscles, and Merkel's discs that differ in their sensitivity to different aspects of mechanical stimulation (*e.g.*, stimulation onset *vs*. sustained touch/pressure), and thus provide complementary information. Similarly to mechanoreceptors, there are multiple types of nociceptors with some responding to high/low temperatures and others to strong mechanical stimulation. Thermoreceptors come in two forms, one type of receptor responds when it is being warmed, and another receptor type responds to cold. Very hot stimuli can sometimes be felt as cold, which is explained by the cold receptors discharging briefly to hot stimuli.

Somatosensory pathways: from the sensory receptors to the cortex

Once the somatosensory stimuli have been converted into action potentials the information reaches, *via* two separate pathways, primary somatosensory cortex at the postcentral gyrus (see **Figure 5-8**). Information from both mechanosensory and nociceptive/temperature receptor neurons enter the spinal cord through the dorsal root ganglion cells. Mechanosensory information ascends *via* a pathway that crosses over to the contralateral side at the level of the medulla and then synapses to the ventral posterior nucleus of the thalamus. Pain and temperature information is mediated by a pathway that crosses over to the contralateral side already at the level of the spinal cord. From the ventral posterior nucleus of the thalamus, information from both of these streams ascend to the somatosensory cortex.

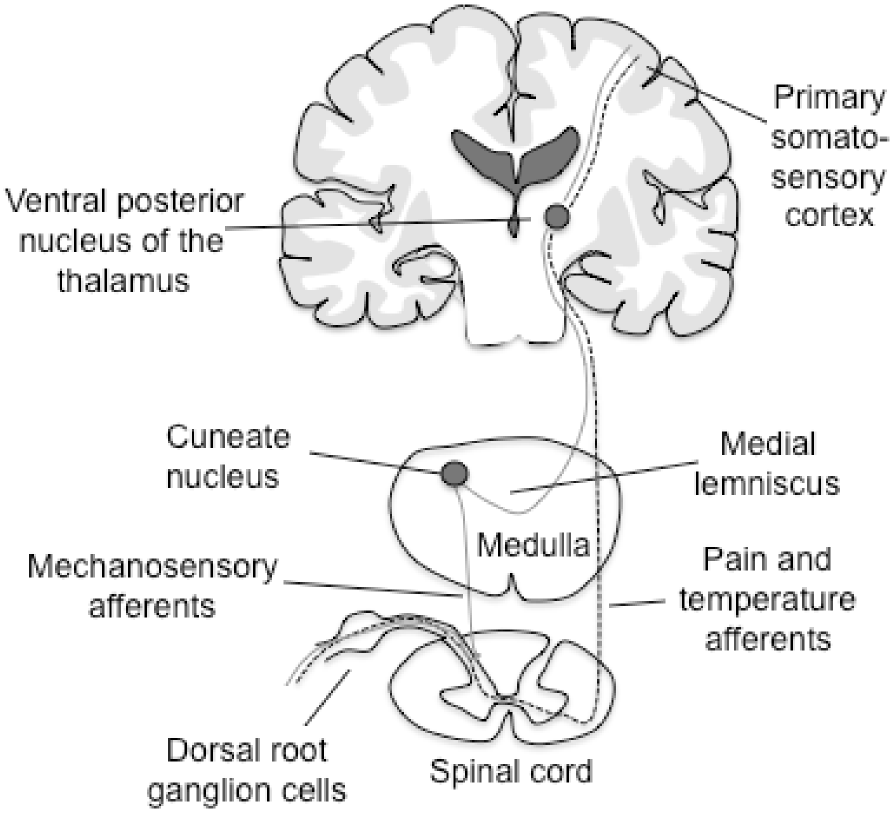


Figure 5-8: Shown are the ascending somatosensory pathways. Both the mechanosensory and pain/temperature afferents ascend through the dorsal root ganglion cells, but whereas the pain and temperature afferents cross over already at the level of the spinal cord, the mechanosensory afferents do not cross over until at the level of medulla / medial lemniscus. All the afferent fibers ascend to the primary somatosensory cortex *via* the ventral posterior nucleus of the thalamus.

Topographic representation of the body: the somatosensory homunculus

Similarly to the retinotopic and cochleotopic representations in the visual and auditory sensory cortices, there is a topographic map of the body in the somatosensory cortex (Penfield and Jasper, 1954). Starting from within the fissure that separates the cerebral hemispheres, where neurons receive information from the genitals, toes, and feet, there is an orderly representation of body parts when progressing to the lateral and inferior directions along the surface of the postcentral gyrus (see **Figure 3-9**). The cortical somatosensory representations of different body parts are disproportional; representation area of the thumb is larger than that of the trunk or that of a leg. This phenomenon is similar to the case of fovea in the visual system, and matches the fine discrimination ability of, for instance, fingers as compared with that of the trunk.

Differences in somatosensory discrimination ability can be measured behaviorally with the so-called two-point discrimination test. In this test, the skin is touched by two pointed/sharp objects (such as the tips of two pencils) and the distance that separates the two points is varied in order to find the distance at which the two points are perceived as separate rather than a single sensation. In the case of fingers, the two points can be very close to one another, whereas on the surface of the leg or trunk the two-point discrimination ability is much less precise.

In addition to the somatotopic representation of the primary somatosensory cortex, there are numerous other somatotopic maps in the secondary somatosensory cortex and other hierarchically higher-order areas. For instance, in premotor cortex in the frontal lobe there are somatotopic maps that are activated when an object is seen in close proximity of a given body part (Fogassi et al., 1996, Brozzoli et al., 2011). These neurons are activated even when there is no actual touch and the responses are most robust when objects are approaching the body. Such representations of peripersonal space possibly underlie the phenomenon of "feeling" an object, or someone else, about to touch one's arm when one sees it.

The olfactory system: sense of smell

The olfactory sense is based on specific chemical receptors in the olfactory epithelium that are located peripherally in the nasal cavity (for a schematic illustration, see **Figure 5-9**). Together these receptors react to a range of different odors and convert the stimulation caused by the odors to electric signals. The action potentials generated by the olfactory receptor neurons are conveyed to so-called mitral cells of the olfactory bulb in the central nervous system. Identification of odors by the brain is believed to be based on distributed patterns of firing that is made possible by combinations of sensitivity profiles of the olfactory receptor neurons (Nara et al., 2011). From the olfactory bulb, there are connections to various brain structures, including the piriform cortex, amygdala, and entorhinal cortex.

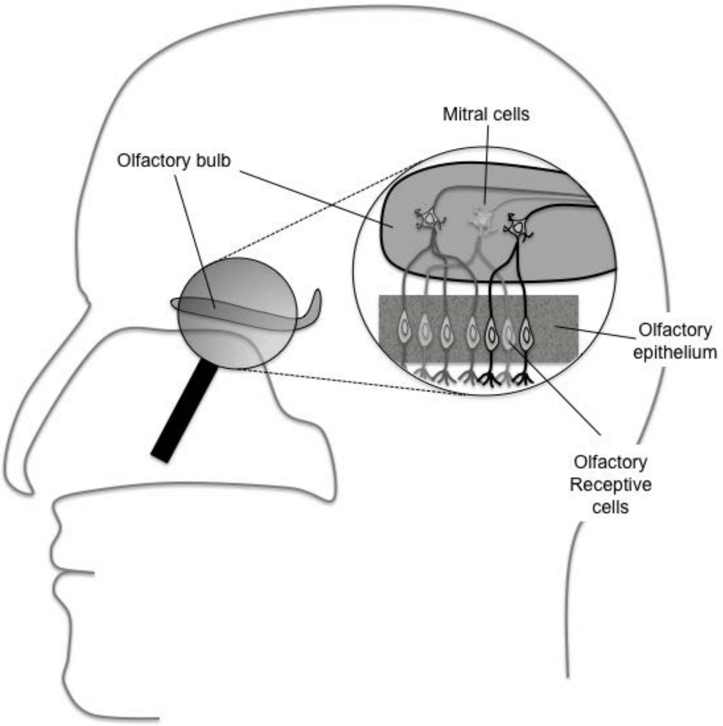


Figure 5-9. Shown is the olfactory system with the olfactory bulb and olfactory receptive cells that reside in the olfactory epithelium. Combination of selective responses by the olfactory receptive cells is thought to underlie the sense of smell.

The sensory systems are inherently multisensory

While the sensory systems have been usually studied separately by presenting visual stimuli in some studies and auditory stimuli in some other studies, it is increasingly recognized that the sensory systems are not as solely devoted to processing of modality-specific information as has been traditionally thought. Rather the sensory systems are to a very high degree multisensory. Interactions between the auditory and visual sensory systems have been studied the most, with visual responses observed in the primary auditory cortex already at very early latencies and *vice versa* (Raij et al., 2010). The multisensory nature of the sensory systems makes it possible to utilize complementary information about perceptual objects provided by visual, auditory, and somatosensory cues.

5.2 The motor system

The motor system is what allows coordinated movements of all sorts: getting up in the morning, preparing breakfast, walking to the car, and driving to work. Most of these movements are fairly automated sequences that have been acquired through practice. Another aspect of motor actions is that most of the movements are part of coherent sequences planned across several timescales. For instance, when we decide to go to work, the subsequent motor behaviors are directed towards attaining that goal even if there are tasks that we face and solve along the way (*e.g.*, overtaking a vehicle in traffic).

Training is another essential aspect of motor coordination. While one cannot probably remember the painstaking process of learning to walk as a toddler, it is possible to remember back to times when one was learning to drive a car. In fact, as will be described later in this chapter, there are partially separate brain mechanisms that are responsible for executing well-learned motor sequences *vs*. novel ones. As an additional aspect of the motor system, it is increasingly realized that the motor system is also involved in perception; understanding how others behave is easier when one has first-hand knowledge oneself on how certain actions are performed. In the following the reader is introduced to these different aspects of the motor system.

There is motor homunculus in the primary motor cortex

Electric stimulation of the primary motor cortex that occupies the precentral sulcus results in involuntary movements of specific muscle groups, depending on the site of stimulation. Similarly to the somatosensory homunculus of the post-central gyrus, there is a motor homunculus in the precentral gyrus, with body parts requiring finer motor coordination, such as fingers and hands, overrepresented (see **Figure 3-9**). For instance, when the thumb representation area is stimulated with transcranial magnetic stimulation, the thumb twitches involuntarily. Descending inputs from the primary motor cortex are conveyed to motor neurons that run from the ventral roots of the spinal cord to muscles and cause contraction of muscle by means of releasing a specific neurotransmitter called acetylcholine.

Frontal cortical hierarchy of motor coordination

There is a posterior-to-anterior hierarchy of motor control in the frontal lobe. Anterior to the precentral gyrus on the lateral aspect of the frontal lobe is the so-called premotor cortex that does not directly innervate specific muscle groups. Rather, the premotor cortex makes it possible to carry out voluntary relatively short planned sequences of movements. This hierarchy is further manifested by the functions carried out by dorsolateral prefrontal cortex, Broca's area that occupies the inferior prefrontal cortical areas, and the supplementary motor cortex that resides in the superior aspects of the frontal cortex and within the inter-hemispheric fissure (see **Figure 5-10**).

The supplementary motor area is involved in planning of sequences of movements, but more based on memory than the premotor cortex that is involved in planning of motor sequences given external stimuli (Shima and Tanji, 1998). Prefrontal cortical areas anterior to the premotor cortex are responsible for hierarchically higher-order action plans, for instance, that one executes sequences of motor movements in an orderly manner when preparing and eating breakfast, driving to work along a route that one has planned, as well as for being able to flexibly modify one's route when there is a traffic accident ahead that has blocked the usual route. Broca's area, in turn, is special in that it is responsible for sequencing speech, as will be descibred in more detail in Chapter 9.

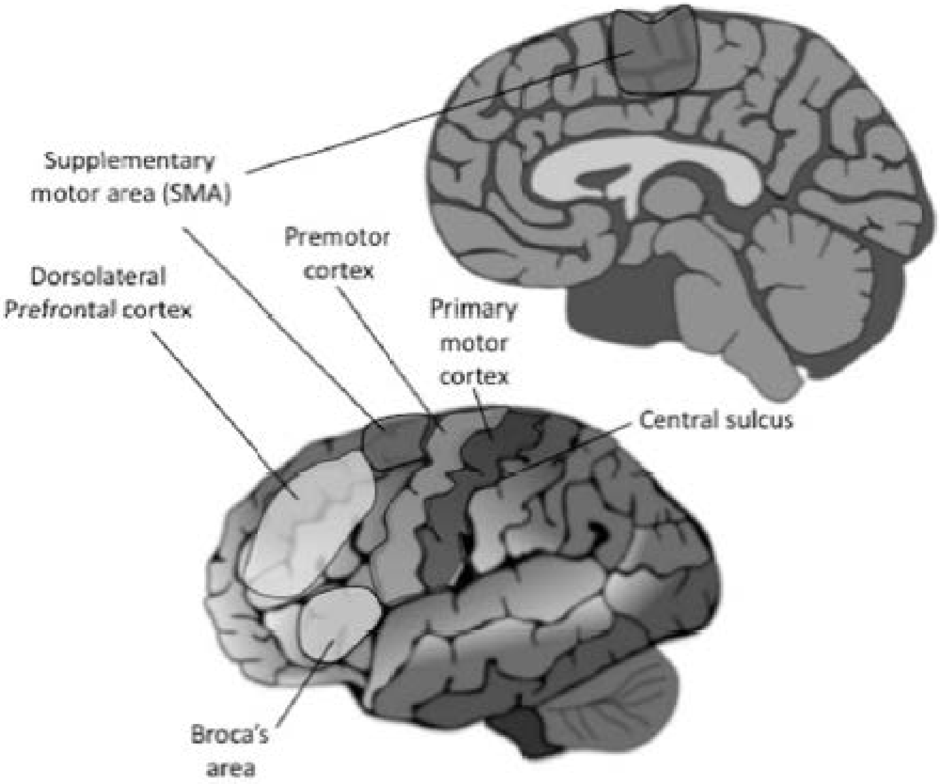


Figure 5-10. Hierarchy of motor control in the frontal cortex. See text for details.

Targeting movements in space: the role of the posterior parietal cortex

The motor areas of the brain are networking with other brain areas when planning and executing motor acts and sequences. One notable example of this is the involvement of posterior parietal cortex in planning of movements. The spatial representations of the posterior parietal cortex make it possible to know where things are in space and, thus, to direct movements in space (for a review, see (Andersen et al., 1997)). In patients with posterior parietal cortex lesions, reaching for objects is disturbed, with patients missing the object even though their reaching movements *per se* are not deficient (*i.e.*, the coordination of the muscle groups itself can be relatively intact, and the reaching act resembles that of a healthy person, but misses the intended target).

Cerebellum and well-learned motor sequences

The cerebellum is a structure that is crucial for fine motor coordination. When a given motor sequence or motor skill (such as ice skating in an ice hockey player or the golf swing in an experienced golfer) has been learned well, following hundreds or thousands of repetitions, the sequence/skill becomes automated and no longer requires voluntary effort / attention to execute. After this type of learning, the cerebellum is mainly responsible for coordinating the motor sequencing.

Basal ganglia and initiation of movements

The basal ganglia are a set of subcortical structures that are critical for facilitating a selected movement / act and suppressing unwanted /competing motor responses. In Parkinson's disease, dopaminergic neurons that innervate basal ganglia are excessively destroyed. Clinically, this cell death is manifested by difficulties in initiating behaviors and in shifting between alternative behaviors. Interestingly, the role of basal ganglia seems not limited to initation of movements, but rather are also involved in selection of and shifting between cognitive sets. For a more detailed description of basal ganglia and their role in goal-directed behavior, see Chapter 10.

Mirror neurons

The motor system is not only involved in execution of movements, but rather also plays a crucial role in perception. The so-called mirror neurons are a good example of this. Originally, mirror neurons were described as neurons that increase their firing rate both when executing an action and when observing the same action performed by others. Rizzolatti and colleagues were the first to describe mirror neurons in the inferior prefrontal cortex of monkeys (for a review, see (Rizzolatti and Graighero, 2004)). Subsequently, the possible roles of mirroring have been extended to also speech/language perception (*i.e.*, the so-called motor theory of speech perception that is described in Chapter 9) and social cognition (*i.e.*, the same neural systems that are activated when experiencing feelings oneself are utilized when inferring the feelings of others, as described in Chapter 12). Overall, this is an important principle that applies both to sensory and motor systems; even though the sensory and motor systems enable perception and acting/moving about, respectively, they are both utilized in other types of cognitive operations such as imagining things and planning future courses of action.

5.3 Concluding remarks

The sensory and motor systems of the brain take up a surprisingly large proportion of the central nervous system. Indeed, perceiving the world and moving about are not trivial tasks, as manifested by difficulties in engineering robots that are able to conduct such tasks smoothly. In fact, some of the processing principles of the brain's sensory and motor systems have been utilized when designing robots. Further, the sensory and motor systems of the brain are not limited to processing of external stimuli and moving the limbs, but rather these systems are constantly utilized in higher cognitive processes, and they to a large extent constitute the "backbone" on which the brain functions. Consequently, knowledge of the fundamentals of sensory and motor systems is highly important for every keen student of cognitive neuroscience. In the next chapter, one of the most central of higher cognitive functions that is intimately tied with the sensory systems will be described, namely attention.

5.4 References

AHVENINEN, J., JAASKELAINEN, I. P., RAIJ, T., BONMASSAR, G., DEVORE, S., HAMALAINEN, M., LEVANEN, S., LIN, F.-H., SAMS, M., SHINN-CUNNINGHAM, B. G., WITZEL, T. & BELLIVEAU, J. W. 2006. Task-modulated "what" and "where" pathways in human auditory cortex. *Proc Natl Acad Sci USA,* 103**,** 14608-14613.

ANDERSEN, R. A., SNYDER, L. H., BRADLEY, D. C. & XING, J. 1997. Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annual Review of Neuroscience,* 20**,** 303-330.

BINDER, J. R., FROST, J. A., HAMMEKE, T. A., BELLGOWAN, P. S., SPRINGER, J. A., KAUFMAN, J. N. & POSSING, E. T. 2000. Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex,* 10**,** 512-528.

BIRAN, I. & COSLETT, H. B. 2003. Visual agnosia. *Curr Neurol Neurosci Rep,* 3**,** 508-512.

BROZZOLI, C., GENTILE, G., PETKOVA, V. I. & EHRSSON, H. H. 2011. fMRI adaptation reveals a cortical mechanism for the coding of space near the hand. *Journal of Neuroscience,* 31**,** 9023-9031.

COX, D. D. & SAVOY, R. L. 2003. Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage,* 19**,** 261-270.

DEYOE, E. A., CARMAN, G. J., BANDETTINI, P., GLICKMAN, S., WIESER, J., COX, R., MILLER, D. & NEITZ, J. 1996. Mapping striate and extrastriate visual areas in human cerebral cortex. *Proc Natl Acad Sci USA,* 93**,** 2382-2386.

FOGASSI, L., GALLASE, V., FADIGA, L., LUPPINO, G., MATELLI, M. & RIZZOLATTI, G. 1996. Coding of peripersonal space in inferior premotor cortex (area F4). *Journal of Neurophysiology,* 76**,** 141-157.

GOODALE, M. A. & MILNER, A. D. 1992. Separate visual pathways for perception and action. *Trends in Neurosciences,* 15**,** 20-25.

HICKOK, G. & POEPPEL, D. 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition,* 92**,** 67-99.

HUBEL, D. H. & WIESEL, T. N. 1959. Receptive fields of single neurons in the cat's striate cortex. *Journal of Physiology,* 148**,** 574-591.

HUBEL, D. H. & WIESEL, T. N. 1968. Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology,* 195**,** 215-243.

ISHAI, A., UNGERLEIDER, L. G., MARTIN, A., SCHOUTEN, J. L. & HAXBY, J. V. 1999. Distributed representation of objects in the human ventral visual pathways. *Proc Natl Acad Sci USA,* 96**,** 9379-9384.

LEWICKI, M. S. 2002. Efficient coding of natural sounds. *Nature Neuroscience,* 5**,** 356-363.

MISHKIN, M. & UNGERLEIDER, L. G. 1982. Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behav Brain Res,* 6**,** 57-77.

NARA, K., SARAIVA, L. R., YE, X. & BUCK, L. B. 2011. A large-scale analysis of odor coding in the olfactory epithelium. *Journal of Neuroscience,* 31**,** 9179-9191.

NELKEN, I., FISHBACK, A., LAS, L., ULANOVSKY, N. & FARKAS, D. 2003. Primary auditory cortex of cats: feature detection or something else? *Biol Cybern,* 89**,** 397-406.

PENFIELD, W. & JASPER, H. 1954. *Epilepsy and the functional anatomy of the human brain,* Oxford, Little, Brown & Co.

RAIJ, T., AHVENINEN, J., LIN, F.-H., WITZEL, T., JAASKELAINEN, I. P., LETHAM, B., ISRAELI, E., SAHYOUN, C., VASIOS, C., STUFFLEBEAM, S., HAMALAINEN, M. & BELLIVEAU, J. W. 2010. Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices. *Eur J Neurosci,* 31**,** 1772-1782.

RAUSCHECKER, J. P. & SCOTT, S. K. 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci,* 12**,** 718-724.

RAUSCHECKER, J. P. & TIAN, B. 2000. Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proc Natl Acad Sci USA,* 97**,** 11800-11806.

RIZZOLATTI, G. & GRAIGHERO, L. 2004. The mirror-neuron system. *Annu Rev Neurosci,* 27**,** 169-192.

SHIMA, K. & TANJI, J. 1998. Both supplementary and presupplementary motor areas are crucial for the temporal organization of multiple movements. *Journal of Neurophysiology,* 80**,** 3247-3260.

STEWART, H., HENDRY, C. & REID, R. C. 2000. The koniocellular pathway in primate vision. *Annual Review of Neuroscience,* 23**,** 127-153.

TALAVAGE, T. M., SERENO, M. I., MELCHER, J. R., LEDDEN, P. J., ROSEN, B. R. & DALE, A. M. 2004. Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity. *J Neurophysiol,* 91**,** 1282-1296.

WEISKRANTZ, L. 1996. Blihdsight revisited. *Curr Opin Neurobiol,* 6**,** 215-220.

ZEKI, S. 1991. Cerebral akinetopsia (visual motion blindness): a review. *Brain,* 114**,** 811-824.

6 Attention

Imagine yourself in a crowded party, where everyone is involved in lively discussions, such as the one illustrated in **Figure 6-1** below; you are involved in one of the many conversations with a couple of colleagues and, amongst all the chatter and noise, you almost literally have to strain your ears to follow what is being said. You may also notice that is easier to follow what your colleague is saying when you look at his/her lip movements. With effort, you can block out the other discussions, until suddenly you are alerted to someone bringing up your name in a nearby conversation.

You might attempt to listen intently and figure out what is being said about you, loosing track of what your colleague is relating at that instant, unless you are able to switch rapidly back and forth between the two discussions to follow them both. Further, as the evening wears on, you might find it increasingly difficult to concentrate on following the conversations, with your tired mind starting to wonder about, for example, reciting on events that took place earlier that day, or about a talk that you are to give the following day. This type of commonplace example from everyday life illustrates phenomena that scientist who study attention and the underlying neural mechanisms focus on.

Figure 6-1: In a noisy cocktail party attention is strained to follow a relevant discussion, yet at the same while one's brain has to also process to some extent the surrounding discussions to alert one, for instance, when someone calls out one's name in the background.

Early definitions of attention

Already in 1890, one of the pioneers of studies of cognitive functions, William James, stated about attention "*Everyone knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others, and is a condition which has a real opposite in the confused, dazed, scatterbrain state*" (James, 1890). Importantly, this early definition by James, which he arrived at mainly through the method of introspection, is still valid on several important points.

The first point that James' definition makes, and one also laid out in the party example above, is that one is constantly faced by the need to select certain information and block out the rest from one's awareness. In modern cognitive neuroscience this phenomenon is called *selective attention* and it is one of the most central research questions in cognitive neuroscience; how is the brain able to filter certain information to awareness while managing to not get distracted by all the other stimuli that constantly bombard one's senses?

In fact, as one indication of the efficacy of selective attention, one typically has to think of situations where there are multiple salient distracters, such as the cocktail party example above, in order to introspectively come to appreciate the challenges faced by the brain during selective attention. In reality, however, there is always infinitely larger amount of both external and internal stimuli than what one can selectively attend. The other significant insight of James' definition is the fact that internal thoughts are often the focus of selective attention in addition to external stimuli. This should be kept in mind when reading through literature on selective attention experiments that typically involve attentional tasks using external stimuli (effects related to internal events would of course be very difficult to measure).

Breakdown of attention in psychiatric and neurologic patients

Certain psychiatric disorders such as schizophrenia are characterized by the inability to focus attention / block out interfering thoughts that arise to the minds of such patients. As a result of this, these patients suffer from extreme confusion and are unable to think or behave in a coherent manner. Blockade of dopamine receptors in striatum alleviate these symptoms in most schizophrenia patients, suggesting that dopamine plays a key role in attentional functions, as will be discussed in more detail later in this chapter.

There are also striking examples of how attention can break down when one suffers damage to the brain. In the so-called unilateral neglect syndrome, which often results from damage to the right parietal (and also frontal) lobe of the brain (Vallar and Perani, 1986), the patient no longer seems to be able to notice anything on the left side of his/her body (see **Figure 6-2**). In conditions called agnosia (for a review, see (Biran and Coslett, 2003)), which most often results from damage to the temporo-parietal junction, patients have difficulty in recognizing and naming perceptual objects. Patients suffering from a so-called associative subtype of agnosia are able to describe what one can do with a phone and are able to draw a phone, but cannot recognize one, including their own drawing of a phone. In a specific type of agnosia called prosopagnosia that results from damage to the fusiform gyrus on the ventral aspect of occipito-temporal junction, patients can no longer recognize faces, even though recognition of other perceptual objects might be preserved (for a review, see (Hecean and Angelergues, 1962)).

While these findings have shed light on the division of attentional mechanisms to those that direct attention in space and those that underlie coherent perceptual object representations, the patient findings are limited given that lesions due to brain damage are rarely circumscribed to one specific brain area and that plastic changes take place during the initial recovery period only after which the assessment of cognitive functions is possible. Thus, compensatory changes might have occurred prior to neuropsychological examination and the deficits might thus not accurately reflect the functional organization in a healthy brain. In the case of psychiatric patients, it is difficult to separate cognitive deficits that are due to significant biological disturbances from those that might result from psychological complications.

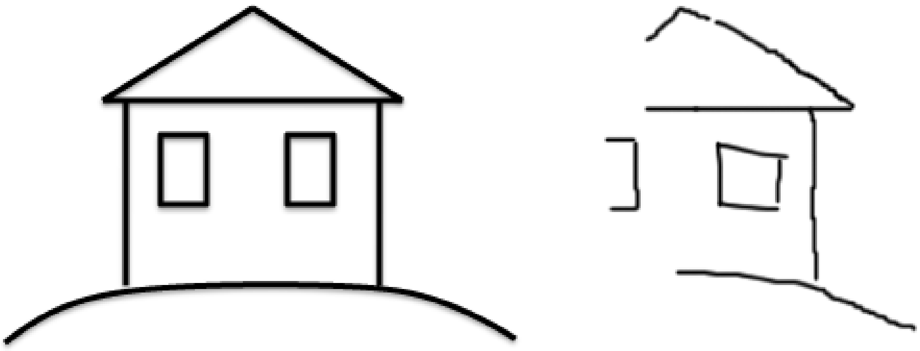


Figure 6-2. Shown is an example of the type of striking manifestations of attentional deficits that patients with unilateral neglect exhibit. When asked to copy a drawing such as the house on the left side of the figure, these patients ignore the left side, leaving it out from the drawing. This deficit is also manifested in daily life in multiple ways, for instance, the patients fail to notice someone talking to them on their left side, write only on the right side of a sheet of paper, and leave the left side of their face unshaved.

Attention is guided by internal models

As will be later discussed in this book in more detail, one always perceives external stimuli in the context of internal models called schemas. These to a great extent shape the way one perceives the world, as proposed already in the 19th Century by pioneers who used the introspection method such as Edward Titchener. In **Figure 6-3**, you can see a concrete example of how such top-down models/representations influence perception. If you have not seen this picture before, it is initially difficult to see what is in the picture, as it looks like pretty much consisting of randomly distributed spots. But once you notice the Dalmatian dog in the picture (depicted on the right side of the figure), it is difficult to look at this picture without the Dalmatian perceptually "popping out" right away. As will be discussed in more detail later, there is a top-down model/representation that guides viewing of the picture, and this shapes the very way the visual cortex processes the incoming stimuli.

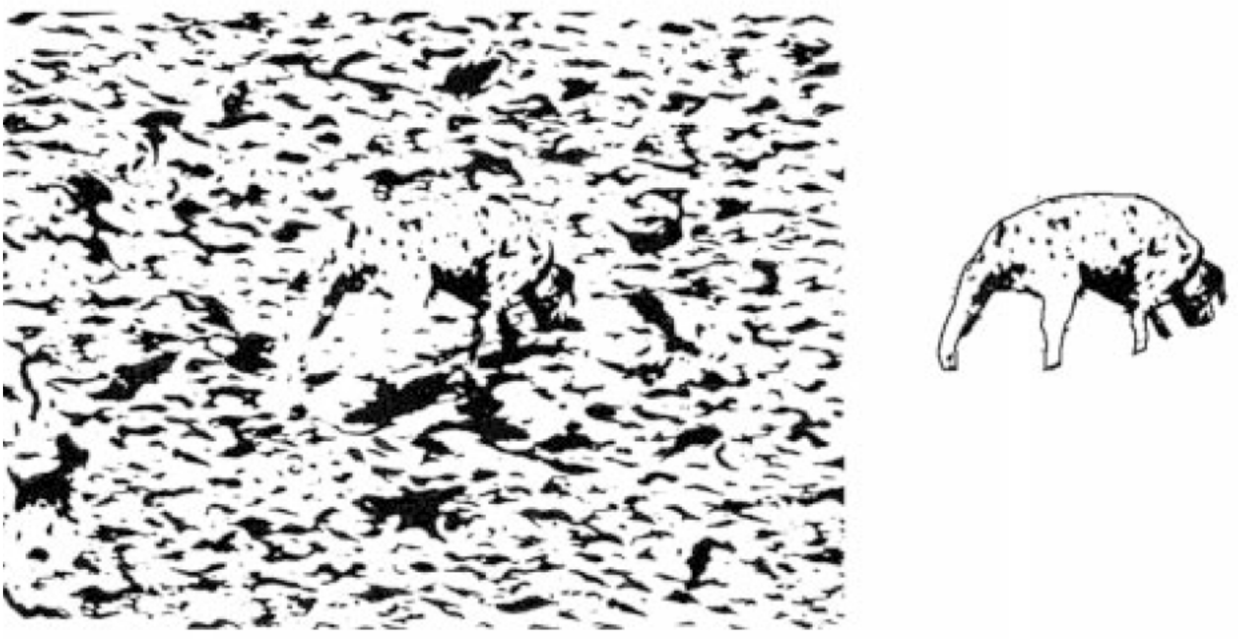


Figure 6-3: Demonstration of how top-down representations guide perception. Left: A Dalmatian dog is camouflaged in this picture, and it takes a while prior to making out the shape of the dog from amongst the spots that make out the surroundings. Right: the shape of the Dalmatian, extracted from the picture shown on the left. Adapted from (Bertrand and Tallon-Baudry, 2000).

Internal models guide eye-movements to gather visual information

The guidance of perception by top-down representations is also manifested in eye movements when one views pictures such as that in **Figure 6-3**. With eye movements, one collects visual information about an object or a visual scene. Eye movements consists of short *fixations*, barely noticeable eye movements called *microsaccades* that occur between the fixations when one is holding gaze on a given detail of interest, large-scale eye movements called *saccades* when eyes are moved around to sample the visual scene, and continuous movement of the eyes called *smooth pursuit* when one is tracking a moving object (for a review on principles governing eye movements and the underlying neural mechanisms, see (Schall and Thompson, 1999)).

Prior to seeing the Dalmatian in the picture above, fixations are relatively evenly distributed across the image, but once the Dalmatian is detected and top-down representation begins to guide perception, fixations concentrate around the edges of the shape of the Dalmatian. In a similar fashion, when viewing pictures of faces, fixations concentrate on the eye and mouth regions, as these are the most informative regions of the face for detecting identity, emotional state, and intentions of others, that are crucial information for smooth social interactions. As another example of the influence of top-down representations on perception and attention, top-down models of semantic context of conversations help predict what is going to be said next (Stephens et al., 2010), thus making it considerably easier to follow conversations during, for example, a noisy cocktail party.

Attention is a multifaceted phenomenon

There are certain aspects of attention that the abovementioned early definition of William James does not include. Since the pioneering days of William James, there has been a plethora of studies on the ability to divide attention between two, or even more, targets (at the time of James, this was referred to as "split attention"). Advances in non-invasive neuroimaging methods has made it possible to study involuntary attention, or in other terms how the brain processes the unattended (Jaaskelainen et al., 2004) and how, for instance, hearing one's own name outside of the focus of attention gains access to one's awareness. Furthermore, distinction has been made between neural mechanisms that shift the focus of attention from one target to another, and those that allow us to sustain attention on a selected target or thought over extended periods of time. This latter ability has also been called vigilance. In the following, each of these attentional functions and the underlying neural mechanisms will be described in detail.

6.1 Selective attention

Law of prior entry

Selective attention, or the ability to focus on a given perceptual object or stream of stimuli over others, has been perhaps the most intensively studied of attentional functions, and today scientists have a fairly good overall understanding of how this task is managed by the brain, even though there still are open questions and issues that need to be clarified. Historically, as already outlined above, William James can be considered as one of the pioneers of selective attention research. Other relevant findings that early researchers arrived at using the method of introspection include the so-called law of prior entry, which was postulated by Edward Titchener (Titchener, 1908).

The law of prior entry refers to a phenomenon where out of two simultaneously presented stimuli, the one that is attended at the time of presentation is perceived as having preceded the unattended one. This behavioral/perceptual level observation suggest that processing of stimuli in the attended "channel" are speeded up and/or those that are not concurrently attended are slowed down, and has inspired subsequent neuroimaging experiments to disclose the underlying neural mechanisms. For instance, in the now classic experiments conducted by Steven Hillyard and colleagues, selectively attended tones elicited the so-called N1 response (that is a correlate of sound detection) in stimulus-trial averaged EEG at an earlier latency than non-attended tones, suggesting that processing of attended sounds is speeded up in the brain (Hillyard et al., 1973).

Behavioral studies of selective attention

Besides the method of introspection that was used in the 19th and early 20th Centuries, and observations of attentional deficits in neurological patients, carefully devised reaction time studies have been used in healthy volunteers to study the principles governing human selective attention. This type of research blossomed especially after transition of research paradigm from behaviorism to cognitive psychology, which made it acceptable to infer "hidden processes" based on changes in reaction times or other behavioral responses as a function of various parametric task and stimulus manipulations. This research approach is also referred to as psychophysics. Behaviorism was the dominant research paradigm until early 1950s, which strictly excluded theoretical/model-based speculation about internal mechanisms, and restricted scientific scrutiny to the objectively measurable. Of course, with the development of neuroimaging methods, it has become possible to measure the inner workings of the brain directly, in addition to quantifying the associated behavioral measures of performance.

Overt and covert attention

Several principles that govern selective attention have been discovered based on inferences drawn from measurements of behavioral responses. One of these principles is the distinction made between *overt* and *covert attention*. In overt attention, the sense organs are turned towards the stimulus that is being attended. In humans, this involves fixing of gaze at the attended object, but monkeys and owls, for instance, also turn their ears towards a sound source. In contrast, covert attention means focusing attention on a part of the visual field or auditory space without moving one's eyes. By instructing subjects to covertly attend an extra-foveal part of the visual field speeds up processing of visual information presented in the attended location (Carrasco and McElree, 2001).

Attentional "spotlight"

The concept of attentional spotlight, specifically, that the focus of attention can be compared to a spotlight illuminating stimuli falling within the attended spatial location, was introduced based on behavioral experiments (LaBerge, 1983). In his pioneering study, LaBerge used a behavioral probe technique in which, during ongoing categorization of five-letter words *vs*. middle letter of the words, the subjects were to respond when the digit "7" (*i.e.*, the probe) occasionally appeared in one of the five letter positions. From the V-shaped function of reaction times across probe positions in the letter categorization task *vs*. relatively flat reaction time function obtained during the word categorization task, it was concluded that the spotlight width was one letter space in the letter categorization task and five letters during the word categorization task.

This finding shows how carefully devised psychophysics studies can reveal significant information about the principles governing human perceptual/cognitive functions and their neural basis. As will be explained later in this chapter in detail, subsequent neurophysiological and neuroimaging studies have shown that the task of the subjects shapes the way neurons in sensory cortices respond to stimuli. These types of changes are then manifested as changes in, for instance, the extent of the attentional spotlight that was documented in the early behavioral studies.

Dichotic listening studies

Dichotic listening experiments, which were motivated by the classic cocktail party example that described in the beginning of this chapter, were set up in the beginning of the 1950s to investigate the degree that unattended messages are processed during selective attention (Cherry, 1953). In a typical dichotic listening experiment that is illustrated in **Figure 6-4**, subjects were presented one audio story to one ear, and another story to the other ear. The task of the subjects was to follow the story to one ear. In some cases the subject was to repeat ("shadow") what was being said.

The questions of interest were how well the subjects could follow the story given the interfering input to the other ear, what they could remember of the unattended passage after the experiment, and if prompted during the experiment, what they could recall of the unattended message. Using various manipulations of the unattended channel, it was noted that "*In no case in which normal human speech was used did the listening subjects fail to identify it as speech; in every such instance they were unable to identify any word or phrase heard in the rejected ear and, furthermore, unable to make definite identification of the language as being English. On the other hand the change of voice – male to female – was nearly always identified, while the 400-cps pure tone was always observed. The reversed speech was identified as having something queer about it by a few listeners, but was thought to be normal speech by others*" (Cherry, 1953).

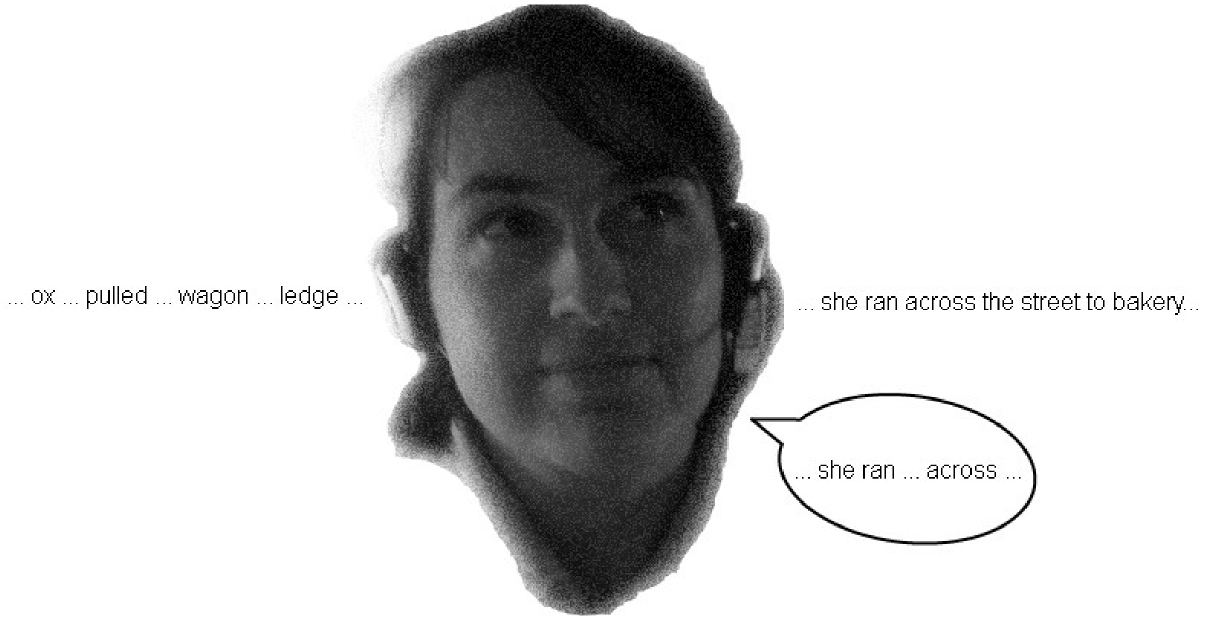


Figure 6-4. In dichotic listening experiments, the task of the subject is to shadow a passage presented to one ear, and ignore the other ear. Occasionally the subjects are prompted to report what they had heard in the unattended ear, to probe the level of processing of the unattended message.

Early and late selection models of selective attention

Based on early work using dichotic listening paradigms, a model was proposed by Donald Broadbent wherein unattended inputs are processed to a very limited degree and only the attended input is processed to semantic level (Broadbent, 1958). Such early selection models of selective attention, however, were soon challenged by findings showing that certain semantic information (*e.g.*, hearing one's own name) occasionally penetrated from the unattended channel to consciousness of the subjects (Treisman, 1969). These findings were interpreted as evidence indicating that stimuli in the unattended channels are processed even up to semantic level, and that they are merely attenuated as compared with the attended input, rather than blocked out. This latter model has been referred to as the late selection model of selective attention (for a comparison of early and late selection models, see **Figure 6-5**). Evidence from modern neurophysiological and functional neuroimaging studies suggest that both the early and late selection models are partly correct.

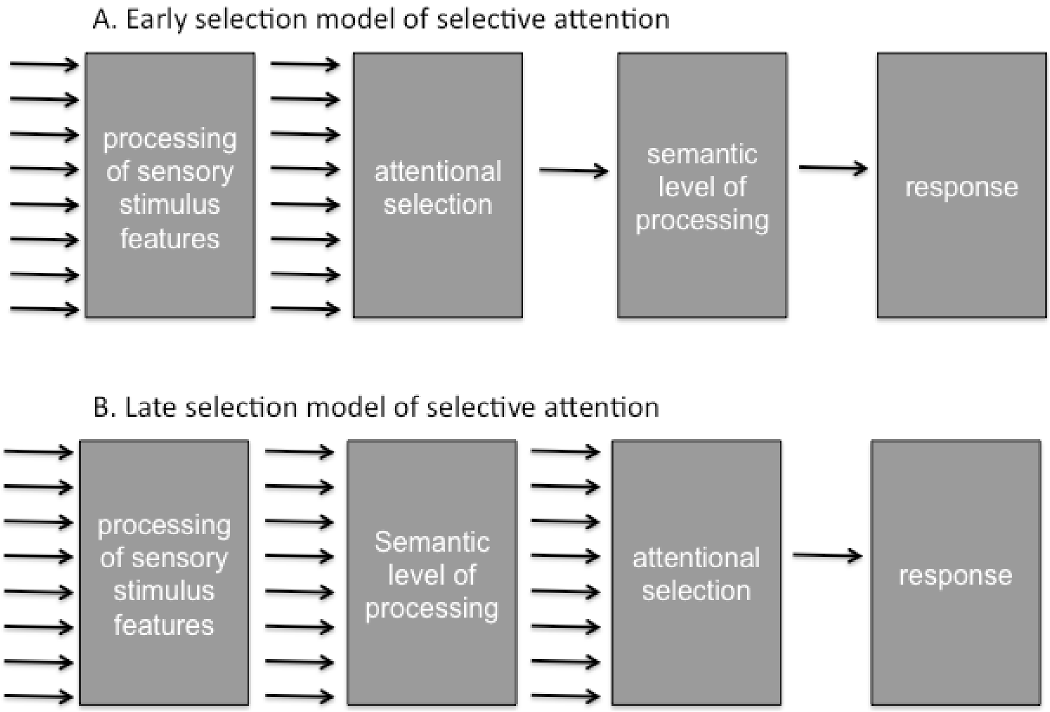


Figure 6-5. Early and late selection models of selective attention. These two models differ in the assumption of the extent to which incoming stimuli are processed prior to attentional selection (*i.e.*, the "bottle-neck" of attention). In the early selection model of attention, inputs from the senses are processed only to the level of sensory feature analysis, whereas in the late selection model, the inputs are processed also semantically prior to attentional selection.

Functional neuroimaging of selective attention

What has been consistently observed in neuroimaging studies is that there is a hierarchy of activations as one progresses from the primary sensory cortices, as was described in Chapter 5; neurons in primary sensory cortices have receptive fields for elementary constituent features of complex auditory or visual objects, such as sounds with relatively simple spectrotemporal features (*e.g.*, a transient ascending sound at a certain center frequency). In secondary auditory areas, the receptive fields are already more complex, responding to combinations of different elementary sound features (Rauschecker, 1998). Higher-order auditory areas are activated by, for instance, speech sounds (Binder et al., 2000).

It has been consistently observed that responses in these brain areas are enhanced to attended stimuli. The higher order areas do, however, respond to stimuli such as speech sounds even when attention is turned away from them. This supports the late selection model of selective attention, as processing of unattended inputs does not stop at the level of early sensory cortices that are viewed as being devoted to processing of elementary sensory features.

Another line of research has attempted to pinpoint the earliest cortical level at which attentional modulations occur. In the most typical experimental setups, scientist have either examined at which anatomical locations selective attention enhances responses of the local neuronal populations and/or at which latencies such modulations occur. Supporting the late selection model of selective attention, there are fMRI studies that found attentional modulation of hemodynamic activity in secondary but not primary auditory cortices (Petkov et al., 2004). However, there are also studies indicating that attentional modulation does occur already in primary sensory cortices and at early latencies (~50 ms from stimulus onset) corresponding to initial processing of the stimuli (Woldorff et al., 1993, Kastner and Ungerleider, 2000, Poghosyan and Ioannides, 2008).

Subcortical contributions to selective attention

Speaking for the early selection model of selective attention, it has been also observed that selective attention can modulate responses in subcortical nuclei such as lateral and middle geniculate of the thalamus and inferior colliculus (Rinne et al., 2008, McAlonan et al., 2008, Jaaskelainen et al., 2011). In fact, anatomical connectivity allows for modulation of stimulus processing at very early stages. There are corticofugal (*i.e.*, from cortex to subcortical structures) connections that are up to ten times more abundant than the ascending connections.

What is noteworthy is that modulatory inputs from the primary auditory cortex can reach cochlear hair cells *via* two synapses, compared with the ascending inputs reaching auditory cortex *via* six synapses (see **Figure 6-6**) (Winer, 2006). As you can recall from Chapter 4, synapses are the tiny gaps between neurons where chemical messaging induces a slight delay to neural transmission compared with the speed of the action potential that traverses down the axon. Thus, top-down efferent inputs can modulate stimulus processing in the auditory system more quickly than the inputs can ascend the system.

How does selective attention enhance processing: gain and tuning

What has been described so far is the level at which attention can influence stimulus processing. But what type of modulations of neuronal processing do the enhanced responses that occur during selective attention reflect? And how does the modulated neuronal processing help pick out the relevant and suppress irrelevant information? This is an important question that the neuroimaging results of enhanced responses during selective attention at various anatomical locations along the hierarchically organized auditory and visual systems do not directly provide an answer to. More specific models of the underlying neural mechanisms have been therefore required, and such models have been recently tested using both neuroimaging in humans as well as sophisticated neurophysiological recordings in experimental animals that have been taught to perform tasks requiring selective attention.

There are currently two models of selective attention with empirical evidence partly supporting both of them. In the so-called gain-enhancement model it is assumed that selective attention works by augmenting concurrently task-relevant and suppressing task-irrelevant stimulus inputs. Vast majority of neuroimaging findings, demonstrating enhanced magnitude for responses to attended stimuli, can indeed be very well explained by increased processing gain for the attended stimuli as compared with unattended stimuli. In fact, the vast majority of neuroimaging studies that have quantified the effects of selective attention have mainly looked at where and at what latencies attention-related signal enhancements can be seen.

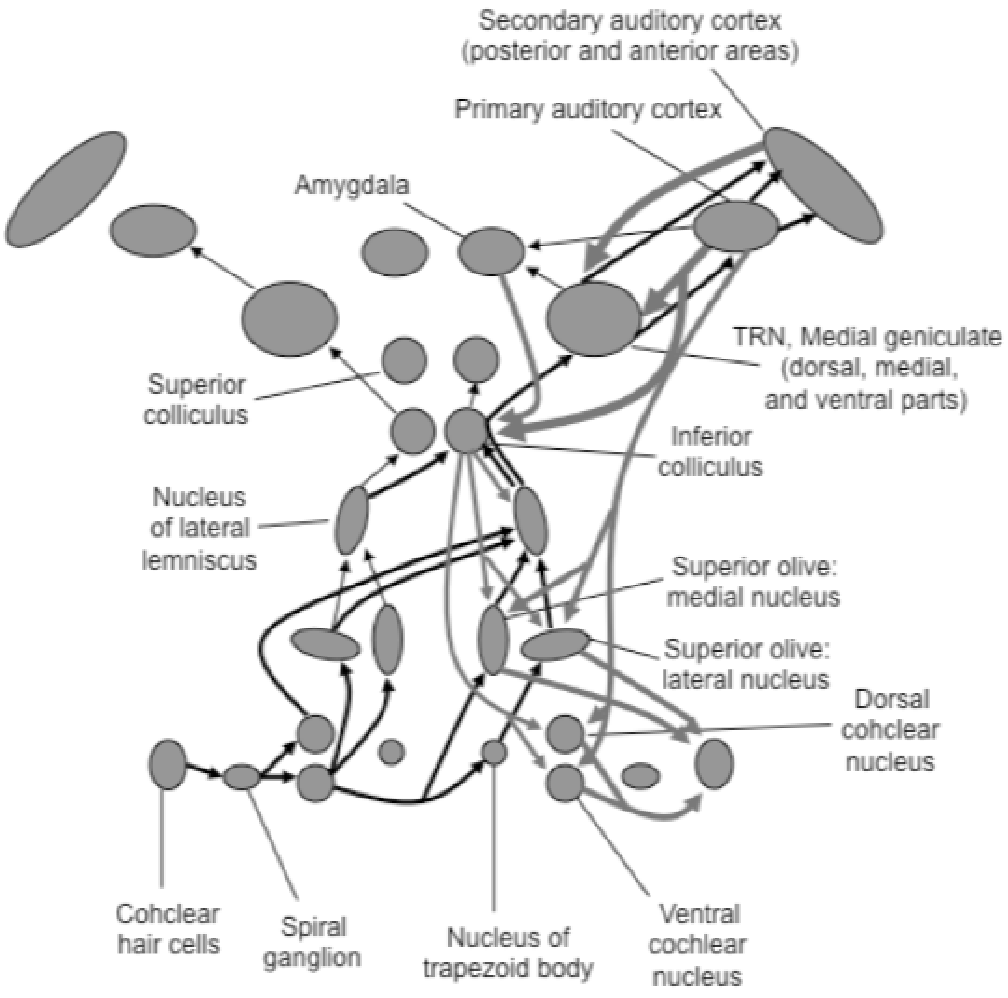


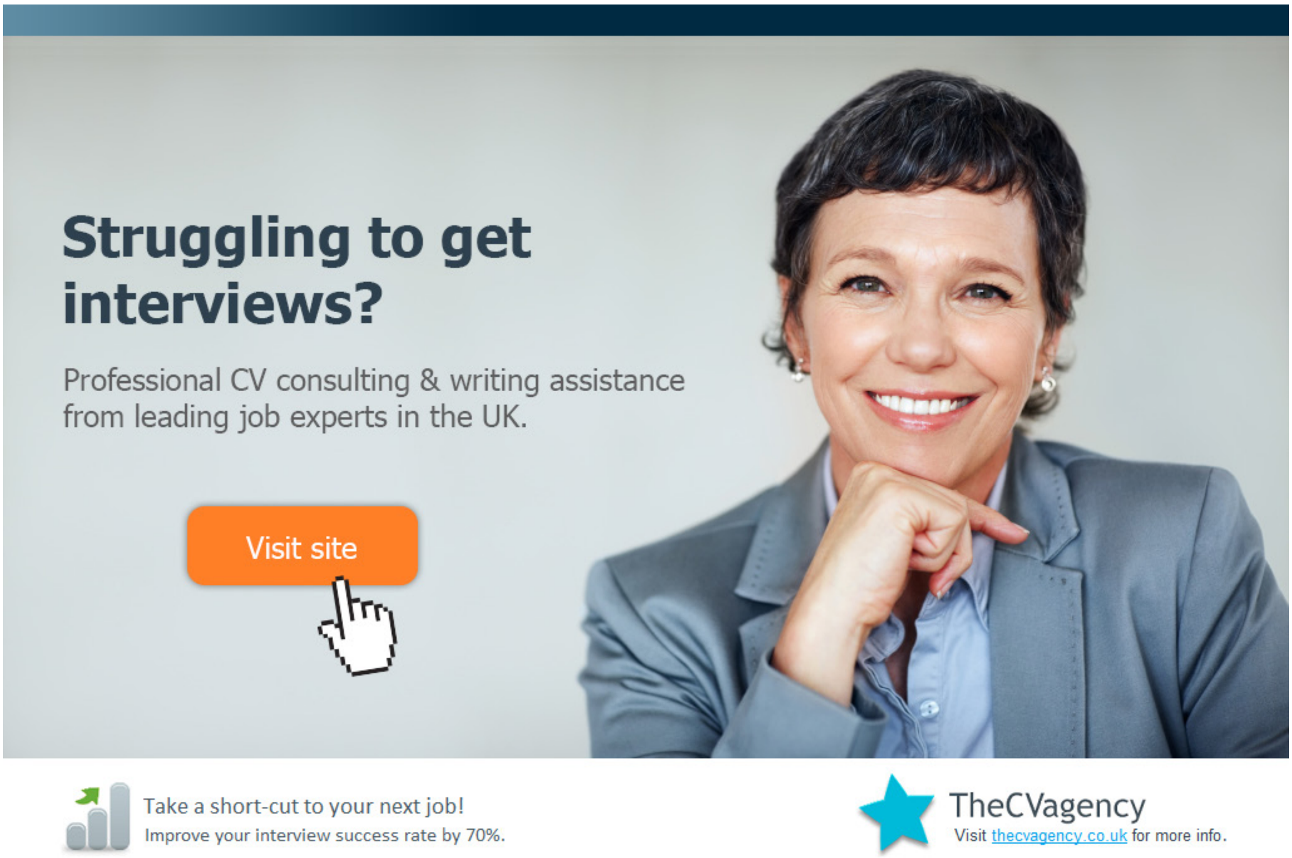
Figure 6-6: Corticofugal connectivity of the auditory system. Here, the corticofugal connections of the auditory system are depicted with thick gray arrows and the ascending connections with black arrows. Note that the corticofugal connections are up to tenfold as many as the ascending connections, are modulatory in nature and, in contrast to the ascending connections, are predominantly ipsilateral (*i.e.*, descend within a given hemisphere instead of crossing over to the other hemisphere). The descending connections from the auditory cortex reach cochlear hair cells *via* only two synapses whereas the ascending inputs traverse from the hair cells to auditory cortex *via* six synapses. This anatomical connectivity supports quick and rich modulation of subcortical processing of sound features, depending upon the task of the subjects.

An interesting alternative model proposes that the very receptive fields of sensory neurons could be shaped or re-tuned to better fit the physical features of the attended object. In a pioneering study conducted by Johathan Fritz and colleagues (Fritz et al., 2003), the spectrotemporal receptive fields of primary auditory cortical neurons were first measured, using specific moving sound patterns called temporally orthogonal ripple combination (TORC) sounds, during a baseline condition. Then, a task was introduced to the animals wherein a tone of a certain frequency indicated access to water, causing the ferrets to attend to the sound frequency of the signal tone.

Spectrotemporal receptive fields are modulated during selective attention

When the TORC sounds were then played to quantify the spectrotemporal receptive field of the neuron, it was observed that the receptive field was modulated to encompass the target sound frequency. This phenomenon only happened if the excitatory center of the receptive field was close enough to the target frequency. These changes occurred relatively quickly and most of them only lingered during task performance. Furthermore, these receptive field changes correlated with behavioral discrimination ability of the animals, which is an important indicator of these modulations being linked with selective attention.

In subsequent studies, similar dynamic receptive field changes were also observed when the target sound consisted of clicks presented at a certain rate and when the target sound was a composite sound involving multiple frequencies. Interestingly, primary auditory cortex neurons were observed to shift their receptive fields to match the features of these target sounds (Fritz et al., 2007). Taken together, findings from these pioneering ferret studies suggest that dynamic receptive field changes occur in primary auditory cortex that act as contrast filters, enhancing the signal-to-noise of the target sounds (see **Figure 6-7**). Subsequently, similar receptive field changes were also observed in monkey visual cortical areas (Womelsdorf et al., 2006), suggesting that this is a common mechanism across the senses, and is not limited to the auditory system.



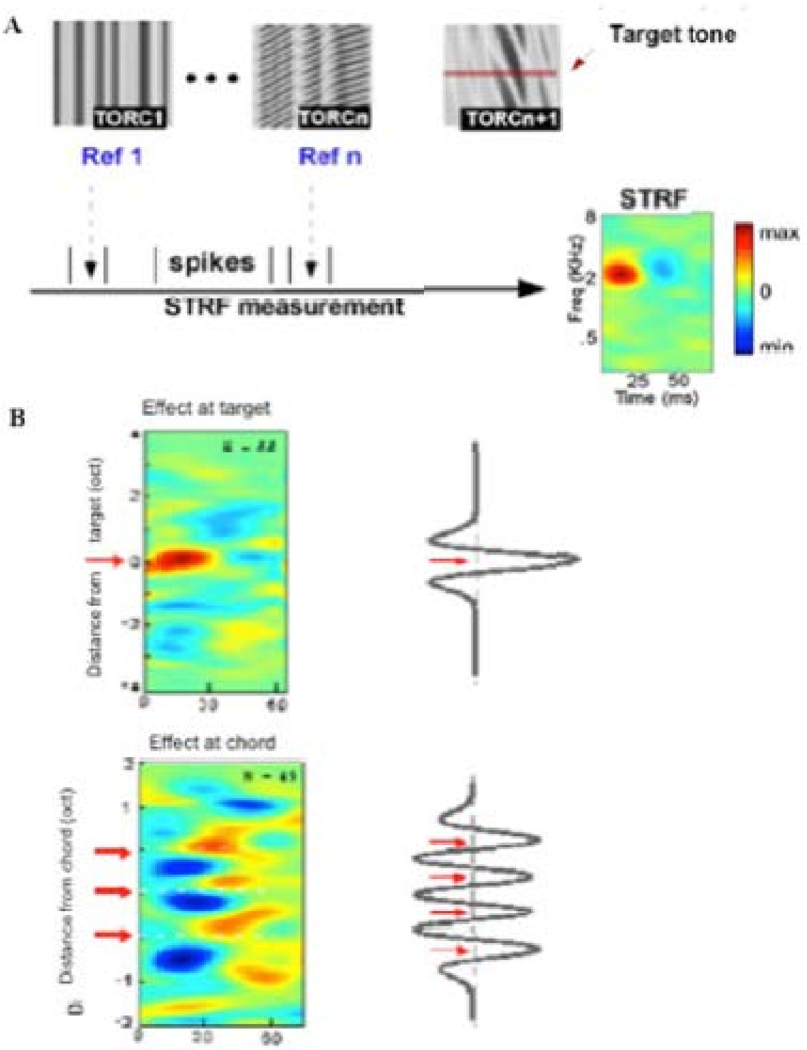


Figure 6-7. Task-specific spectrotemporal receptive field changes in the primary auditory cortex of a ferret. A) Receptive fields were quantified by presenting temporally orthogonal ripple combination (TORC) stimuli, broadband noise superpositioned by envelopes of ripples that drifted up and down, to the ferrets. Neural responses to these TORCs were then cross-correlated with the TORC spectrograms to reveal the receptive field of the neuron. B) On the left are shown difference receptive fields depicting dynamic changes in receptive fields with a single target tone (upper part) and when the animals were attending a multiple-tone chord (lower part). On the right are shown task-related signature profiles of auditory-cortex effects that resemble contrast-match filters. Adapted from (A) (Atiani et al., 2011) and (B) (Fritz et al., 2007).

These findings support the early selection model of selective attention in the sense that the changes occurred already at the level of primary auditory cortex. Further, these findings shed light on how selectively attended stimulus features are picked out from amongst the noise: the number of neurons responding to the features of the attended stimuli increases due to retuning of receptive fields, thus increasing the signal-to-noise ratio. While there are also findings suggesting that selective attention merely modulates the gain of stimulus processing, by enhancing responses to attended inputs and suppressing responses to unattended inputs, recent work suggests that the retuning effects would be seen under circumstances where the to-be attended and unattended stimuli are partially overlapping the same neuronal receptive fields (Lee and Maunsell, 2009). There is yet limited evidence of retuning effects in humans, however, recent non-invasive neuroimaging studies have suggested that retuning does occur in a matter of seconds in human secondary auditory cortical areas after shifting the direction of selective attention (Ahveninen et al., 2011). Furthermore, these retuning effects predicted the accuracy of attentional performance, which is again a key finding that links the retuning effects to the ability to attend selectively.

6.2 Involuntary attention

Keeping the guard on the unattended

While the findings described above provide valuable insight into how selective attention facilitates perception, such as being able to comprehend what our conversation partner is saying in a noisy cocktail party even though we can barely hear him/her above the chatter, these findings do not *per se* describe how the focus of attention is swiftly captured from the ongoing task when something unexpected happens outside of the focus of attention, such as someone mentioning our name in a background conversation. In the early days of attention research this phenomenon was called passive attention. The modern term for this phenomenon is involuntary attention.

During selective filtering of task-relevant information, the brain faces the challenge of having to keep a guard on for the occurrence of potentially significant stimuli or events outside of the focus of attention. In the pioneering dichotic listening experiments described above, intrusion of the unattended was attained by inserting personally significant sounds, such as the name of the subject, to the unattended channel. Everyone can also probably remember occasions where something moving in the periphery of ones vision, or an unexpected sudden sound, resulted in re-orienting of attention.

In cognitive neuroscience, it has been an important topic of investigation how the brain determines novelty of stimuli outside of the focus of attention, and how attention is captured by novel stimuli. Processing of novelty in terms of physical stimulus features has been further dissociated from the processing of stimulus significance, as in the case of processing of emotion-laden personally significant stimuli. As an example of the latter, shorter-latency of early EEG responses have been documented to personally relevant sounds (one's own mobile phone ringing sound) in auditory and frontal areas (Roye et al., 2010).

Findings from oddball paradigms

In auditory novelty processing experiments, subjects have been typically presented with various types of sounds either in silence, or when interspersed as physically deviant stimuli between repetitive streams of homogeneous sounds called "standard stimuli". This latter type of experimental paradigm is called an oddball paradigm. What has been consistently observed under these experimental conditions is that presentation of a sound after a long silence elicits a robust response originating from the auditory cortices. This response has been called the N100 response given its characteristic latency of occurrence of about 100 ms from sound onset (for a review see (Näätänen and Picton, 1987)).

The response to novel sounds when they are interspersed by standard sounds that they deviate from is typically delayed in latency and suppressed in amplitude (Naatanen et al., 1978). Behaviorally, the magnitude of this response correlates with the extent that the novel sounds distract ongoing auditory and visual tasks, thus linking these responses to more direct measures of attentional capture. There are currently alternative theoretical accounts on whether the delayed and amplitude-suppressed response reflects activation of an independent set of neurons that keep track of auditory regularities (Naatanen et al., 2005), or whether the delayed and suppressed response is explained by short-lived adaptation of feature-specific neurons that give rise to the response that is observed when a sound is presented after a long silence (Jaaskelainen et al., 2004, May and Tiitinen, 2009).

Neuronal adaptation as a mechanism that filters stimulus novelty

As you can recall from Chapter 5, one of the hallmarks of sensory systems is the organization of stimulus processing into parallel and hierarchical processing streams; in subcortical structures and primary sensory receptive areas neurons respond to more elementary stimulus features than in secondary and higher order cortical areas. In the visual cortex, the classic findings of Hubel and Wiesel laid out a scheme in which there are simple cells responding to a single orientation of a contrast or line, complex cells responding to moving lines of a given orientation and hypercomplex cells responding to an edge moving in a certain direction (Hubel and Wiesel, 1968).

In the auditory modality, cells responding to a certain sound frequency can be found in the primary auditory cortex, whereas neurons in secondary auditory cortical areas show complex receptive fields with multiple frequencies and/or sensitivity to more complex spectrotemporal sound components (Rauschecker, 1998). The division of processing to parallel processing of object identity ("*what*") and location of objects in space ("*where*") also seems to be a common organizing principle across the different sense modalities.

While this organization neatly explains how the identities and locations of perceptual objects are decoded quickly and efficiently by the brain, it cannot alone explain how stimulus novelty is assessed. An additional principle of especially cortical neurons is that they, once stimulated, are adapted for a short duration of time. During this time, if a similar stimulus is presented, the response of cortical neurons that are sensitive to it yield a much diminished response. Given that this diminution of cortical responses co-occurs with diminished extent to which the stimuli capture one's attention, it seems that cortical stimulus-specific adaptation offers a relatively simple and straightforward mechanism that could determine stimulus novelty in the brain (for a review, see (Jaaskelainen et al., 2007)). For a schematic illustration showing this processing principle, see **Figure 6-8**.

Indeed, if you think of a cocktail party or other noisy environment, most of the sounds are repeating such as the characteristic voices of speakers. Then, one might notice a new speaker joining a conversation, or the sound of a glass breaking, as it makes a sound that is physically distinct amongst the chatter. There is also evidence suggesting that the dorsal ("*where*") pathway is responsible for this type of fast and coarse initial analysis of stimulus novelty in both the visual (Bar et al., 2006) and auditory (Jaaskelainen et al., 2004) sense modalities.

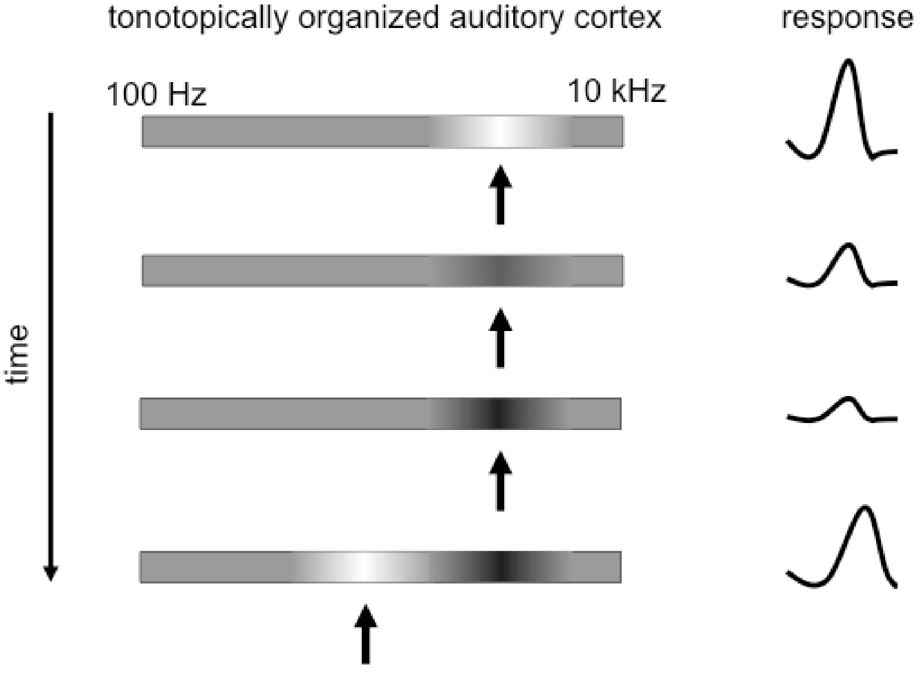


Figure 6-8. A schematic illustration showing how neuronal short-lived adaptation can enable gating of novelty as happens when novel sounds occur amongst unattended sounds, such as a glass breaking among the chatter of a cocktail party. Here, the case of novelty discrimination in sound frequency is illustrated by showing what happens as a function of time when a sound (indicated by topmost of the black arrows) excites tonotopically organized auditory cortex. Since there are no preceding sounds, this sound elicits a robust response. After the response the patch of tonotopic cortex that was excited is transiently (~seconds) inhibited. Upon repeating the same sound frequency, indicated by the two arrows in the middle, the response is diminished due to the lingering inhibition. When a sound of a different frequency, indicated by the lowermost arrow, excites only partially inhibited part of the tonotopically organized cortex, a robust response is elicited and the sound captures attention. Note that the response is slightly delayed and suppressed given that there is in this example some lateral inhibition, caused by the preceding sounds, present in the part of the tonotopically organized patch that the fourth sound excites.

Emotionally associated sounds are screened by amygdala

Involuntary attention shifts to novel sounds is not, however, solely explained by cortical processing. Personally significant stimuli, such as hearing one's own name or the voice of a loved one, are pronouncedly effective in capturing attention. As will be described in more detail in Chapter 11, there are direct connections from the sensory nuclei of the thalamus, such as the auditory middle geniculate body, to amygdala. Coding of emotionally significant stimuli in the amygdala forms a mechanism through which sounds with emotional associations can robustly capture attention from ongoing tasks.

Orienting response

The involuntary shift of attention, especially to emotion-laden stimuli, is often associated with overt orienting and autonomic nervous system responses such as increase in heart rate and blood pressure. This effect has been termed orienting response and it was first described by Russian scientists (Sokolov, 1963). The orienting response helps organisms to prepare to either flee or fight when faced with unexpected stimuli that signal, for instance, the approach of threatening predators. Orienting response does not, however, occur every time that novel stimuli catch one's attention.

6.3 Voluntary shifting and dividing of the focus of attention

Besides the questions of how the brain filters selectively attended information from noise and how attentional focus is captured by significant events outside the attentional focus, it has been one of the central questions of cognitive neuroscience how we voluntarily choose between alternative attentional targets, and what is the cascade of brain processes that are associated with shifting of attentional focus. Further, as mentioned above this question also relates to the problem of how one is able to divide attention between two or more targets, as this involves quick shifting of attentional focus back and forth between two or more perceptual objects.

Cerebral events associated with cue-triggered voluntary attention shift

Neuroimaging experiments where attentional shifting has been investigated have mostly used cued paradigms, where an external stimulus such as arrow pointing to left *vs.* right serves as a cue to shift the focus of covert attention (Posner, 1980). This is often referred to as the Posner paradigm after the developer of the paradigm, Michael Posner. In neuroimaging studies that have used this type of paradigm, it has been observed that dorsolateral prefrontal cortical areas, together with superior posterior parietal areas show transient activity when an attention shift is commenced (see **Figure 6-9**). It can be of course argued to what extent the results obtained in such cued attention shift paradigms are representative of the neural events underlying a genuinely voluntary attention shift as the attention shifts in the cued paradigms are triggered by external stimuli.

Whether the parietal or frontal areas are the more central structure for a voluntary attention shift has been a topic where there are findings supporting both possibilities. Further, a hierarchy has been observed in parietal cortex where more superior aspects of the posterior parietal cortex show transient activity after the cue stimuli and areas slightly inferior to these show sustained activity that follows the transient response (Yantis et al., 2002). This sustained activity is probably associated with maintenance of the attentional focus between the attention shifts. Furthermore, it can be assumed that the selective-attention induced retuning of neuronal receptive fields that occurs in sensory cortical areas (Fritz et al., 2003, Ahveninen et al., 2006) is triggered and maintained by the prefrontal/posterior parietal network (for animal findings suggesting prefrontal cortex involvement, see (Fritz et al., 2010)).

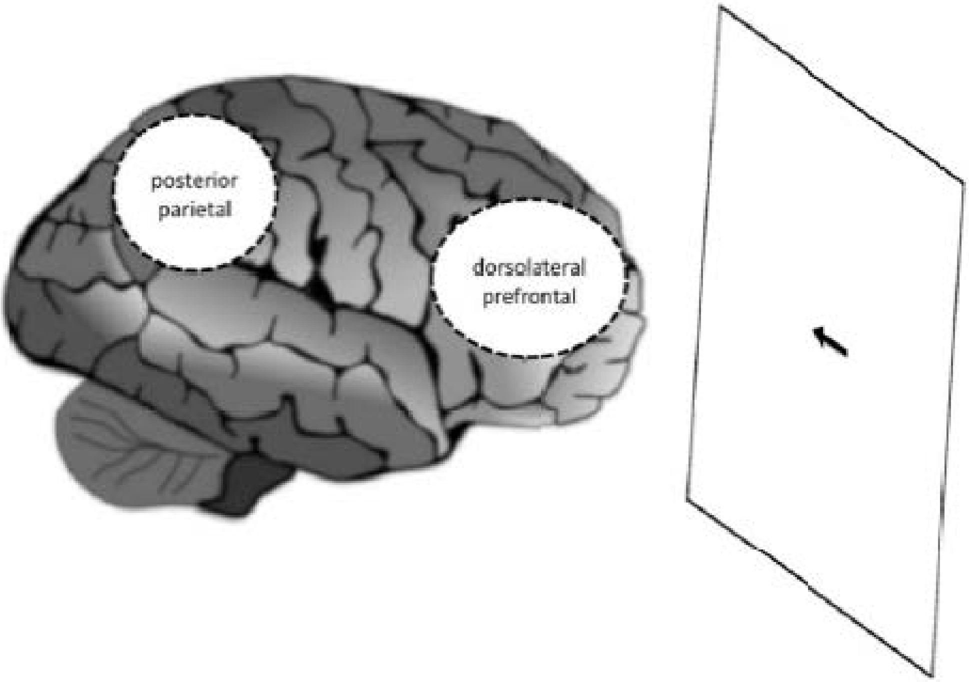


Figure 6-9: Cortical areas involved in voluntary shifting of attentional focus. In a typical experimental paradigm, the direction of the arrow on the screen changes occasionally and signals the subjects to shift focus of attention to left *vs*. right hemifield or, in the case of auditory studies, to left *vs*. right ear. The posterior parietal areas, consisting of superior parietal lobule (SPL) and inferior parietal sulcus (IPS), have been most often implicated in this type of voluntary re-directing of attention. The posterior parietal cortex areas seem to exhibit hierarchy with right SPL triggering transient attention shifts and left IPS maintaining the attentional focus between the attention shifts (Yantis et al., 2002). There are also other areas involved, such as the frontal eye fields residing in the superior aspects of posterior frontal cortex that direct eye movements and sensory cortices where dynamic gain and receptive field changes take place to help filter task-relevant stimuli from task-irrelevant ones.

Basal ganglia are involved in shifting between alternative actions

Subcortical structures also play a role in directing attention, as initiating behavior and switching from one behavioral plan to another is regulated by basal ganglia. This process is presented in detail in Chapter 10. In short, it is known that signals from prefrontal cortex reach the basal ganglia through the so-called direct pathway that leads to release from inhibition of selected action, and induces inhibition of competing inputs through another pathway called the indirect pathway, at the level of thalamus. While the role of basal ganglia has been mostly inspected in the context of regulation of motor actions, it seems that the motor system is overall involved in setting goals and implementing action plans.

In Parkinson patients, destruction of dopaminergic neurons that ascend from pars compacta region of the substantia nigra to striatum results in inability to initiate movements. While this disorder was for a long time predominantly considered as a motor disease, these patients also suffer from deficits in shifting from one cognitive set to another (Bowen et al., 1975, Cools et al., 2001). Thus, the behavioral plans form hierarchically higher order "sets" that guide how the focus of attention is shifted from one target to another while one navigates through the tasks of daily life. In addition to dopamine, there are other chemical neurotransmitter substances that contribute to attention. These will be described next.

6.4 Neurotransmitter basis of attention

While all brain function is ultimately based on neurochemistry (as there would not be neuronal signaling and thus brain activity without chemical neurotransmitter substances) certain neurotransmitters play a larger role than others in specific attentional functions. As has been already described above, dopamine has been linked to selective attention. In schizophrenia, one of the characteristic deficits is an inability to block out irrelevant external stimuli and thoughts. These problems are alleviated by dopamine receptor antagonists such as haloperidol that bind to dopamine receptors without activating them, thus deterring dopamine from exerting its effects on neural function. Conversely, patients suffering from Parkinson's disease are known to have difficulties shifting attentional focus from one task to another. Similar "sticky cognition" has been observed in healthy volunteers who have been administered dopamine D2-receptor antagonist (Mehta et al., 1999). Based on observations such as these, it has been concluded that dopamine is intimately involved in regulating selective attention.

Serotonin and norepinephrine regulate cortical arousal

5-hydroxytryptamine, also known as serotonin, and norepinephrine also known as noradrenaline, play a key role in regulating the level of cortical arousal, vigilance and sensory amplification. In early studies conducted in cats, it was noted that brainstem reticular activation system was important for regulation of sleep and wakefulness (Moruzzi and Magoun, 1949). Subsequent research has shown that the reticular activation system can be divided into distinct nuclei with differential neurotransmitter mechanisms; serotonergic inputs to cortex mainly originate in brain stem raphe nuclei whereas norepinephrinergic inputs ascend from locus ceruleus (for a recent review of these systems and mechanisms of sleep, see (Saper et al., 2005)). Low levels of especially serotonin have been shown in clinical depression that is characterized by overall lack of energy and inattention.

The major dopaminergic, serotonergic and norepinephrinergic tracts of the brain are depicted in **Figure 6-10**. Note that there are a number of different types of dopamine, serotonin and norepinephrine receptors and thus the effects that these neurotransmitters have on post-synaptic neurons vary (for a review, see (Cooper et al., 2003)). As a result of this, the effects of for a given neurotransmitter can be either depolarizing or hyperpolarizing, and the effect can be fast, when the receptors directly gate ion channels, or relatively slow and modulatory when receptor binding results in second messenger mediated opening or closing of ion channels (see Chapter 4).

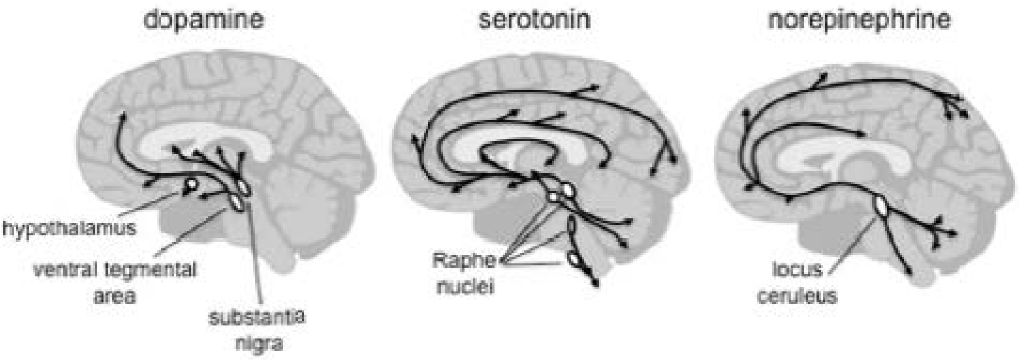


Figure 6-10: Shown are the major dopaminergic, serotonergic and norepinephrinergic pathways of the human brain.

Inhibitory transmitters and attention

In addition to dopamine, serotonin, and norepinephrine, inhibitory neurotransmitters are important for attentional functions. As described above, post-stimulus lateral inhibition is one of the candidate mechanisms underlying involuntary attention. Lateral inhibition is also a candidate mechanism that enhances stimulus discrimination / retunes neural receptive fields during selective attention (for a review, see (Jaaskelainen et al., 2011)). Inhibitory interneurons that exert their effects by releasing inhibitory amino acid called gamma-aminobutyric acid (GABA) are thus important for selective attention. GABA, when reaching the post-synaptic membrane, works mainly by binding to $GABA_A$ receptors that lead to opening of Cl^- channels. This leads to influx of Cl^- and hyperpolarizes the neuron, thus diminishing the firing rate of the neuron (see Chapter 4).

6.5 Concluding remarks

The ability to attend is one of the most important of human cognitive functions. It makes it possible for one to direct limited cognitive resources to concurrently relevant aspects of the environment and/or mental processes such as memorizing critical facts pertaining to the situation at hand. In several pathological conditions, deficits of attention handicap the patient from performing even the simplest of everyday activities. At the neural level, selective attention to certain stimuli causes re-tuning of neural receptive fields at the level of sensory cortices that helps filter attended stimuli to awareness. Prefrontal and posterior parietal cortical mechanisms, in turn, seem to be most relevant for directing of attention to relevant aspects of the environment and, thus, in inducing the receptive field shifts. Mechanisms of neuronal adaptation, on the other hand, appear to underlie novelty detection among the unattended stimuli. Several neurotransmitter systems are important for intact attentional functions, including dopamine, serotonin, and norepinephrine. In the next chapter, memory functions will be reviewed that are highly relevant for selective attention; after all, it is what one remembers (*i.e.*, knowledge of the world and for instance personalities of others) that allows one to direct attention to the concurrently most relevant aspects of the environment.

6.6 References

AHVENINEN, J., HÄMÄLÄINEN, M., JAASKELAINEN, I. P., AHLFORS, S. P., HUANG, S., LIN, F.-H., RAIJ, T., SAMS, M., VASIOS, C. & BELLIVEAU, J. W. 2011. Attention-driven auditory cortex short-term plasticity helps segregate relevant sounds from noise. *Proc Natl Acad Sci USA,* 108**,** 4182-4187.

AHVENINEN, J., JAASKELAINEN, I. P., RAIJ, T., BONMASSAR, G., DEVORE, S., HAMALAINEN, M., LEVANEN, S., LIN, F.-H., SAMS, M., SHINN-CUNNINGHAM, B. G., WITZEL, T. & BELLIVEAU, J. W. 2006. Task-modulated "what" and "where" pathways in human auditory cortex. *Proc Natl Acad Sci USA,* 103**,** 14608-14613.

ATIANI, S., ELHILALI, M., DAVID, S. V., FRITZ, J. B. & SHAMMA, S. A. 2011. Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. Neuron, 61, 467-480.

BAR, M., KASSAM, K. S., GHUMAN, A. S., BOSHYAN, J., SCHMID, A. M., DALE, A. M., HÄMÄLÄINEN, M. S., MARINKOVIC, K., SCHACTER, D. L., ROSEN, B. R. & HALGREN, E. 2006. Top-down facilitation of visual recognition. *Proc Natl Acad Sci USA,* 103**,** 449-454.

BERTRAND, O. & TALLON-BAUDRY, C. 2000. Oscillatory gamma activity in humans: a possible role for object representation. *International Journal of Psychophysiology,* 38**,** 211-223.

BINDER, J. R., FROST, J. A., HAMMEKE, T. A., BELLGOWAN, P. S., SPRINGER, J. A., KAUFMAN, J. N. & POSSING, E. T. 2000. Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex,* 10**,** 512-528.

BIRAN, I. & COSLETT, H. B. 2003. Visual agnosia. *Curr Neurol Neurosci Rep,* 3**,** 508-512.

BOWEN, F. P., KAMIENNY, R. S., BURNS, M. M. & YAHR, M. 1975. Parkinsonism: effects of levodopa treatment on concept formation. *Neurology,* 25**,** 701-704.

BROADBENT, D. A. 1958. *Perception and communication,* London, Pergamon Press.

CARRASCO, M. & MCELREE, B. 2001. Covert attention accelerates the rate of visual information processing. *Proc Natl Acad Sci USA,* 98**,** 5363-5367.

CHERRY, E. C. 1953. Some experiments on the recognition of speech, with one and with two ears. *Journal of the acoustical society of America,* 25**,** 975-979.

COOLS, R., BARKER, R. A., SAHAKIAN, B. J. & ROBBINS, T. W. 2001. Mechanisms of cognitive set flexibility in Parkinson's disease. *Brain,* 124**,** 2503-2512.

COOPER, J. R., BLOOM, F. E. & ROTH, R. H. 2003. *The biochemical basis of neuropharmacology,* New York, Oxford University Press.

FRITZ, J., ELHILALI, M., DAVID, S. V. & SHAMMA, S. A. 2007. Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1? *Hearing Research,* 229**,** 186-203.

FRITZ, J., SHAMMA, S., ELHILALI, M. & KLEIN, D. 2003. Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat Neurosci,* 6**,** 1216-1223.

FRITZ, J. B., DAVID, S. V., RADTKE-SCHULLER, S., YIN, P. & SHAMMA, S. A. 2010. Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal cortex. *Nature Neuroscience,* 13**,** 1011-1019.

HECEAN, H. & ANGELERGUES, R. 1962. Agnosia for faces (prosopagnosia). *Arch Neurol,* 7**,** 92-100.

HILLYARD, S. A., HINK, R. F., SCHWENT, V. L. & PICTON, T. W. 1973. Electrical signs of selective attention in the human brain. *Science,* 182**,** 177-180.

HUBEL, D. H. & WIESEL, T. N. 1968. Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology,* 195**,** 215-243.

JAASKELAINEN, I. P., AHVENINEN, J., ANDERMANN, M. L., BELLIVEAU, J. W., RAIJ, T. & SAMS, M. 2011. Short-term plasticity as a neural mechanism supporting memory and attentional functions. *Brain Research,* 1422**,** 66-81.

JAASKELAINEN, I. P., AHVENINEN, J., BELLIVEAU, J. W., RAIJ, T. & SAMS, M. 2007. Short-term plasticity in auditory cognition. *Trends in Neurosciences,* 30**,** 653-661.

JAASKELAINEN, I. P., AHVENINEN, J., BONMASSAR, G., DALE, A. M., ILMONIEMI, R. J., LEVANEN, S., LIN, F.-H., MAY, P., MELCHER, J., STUFFLEBEAM, S., TIITINEN, H. & BELLIVEAU, J. W. 2004. Human posterior auditory cortex gates novel sounds to consciousness. *Proc Natl Acad Sci USA,* 101**,** 6809-6814.

JAMES, W. 1890. *The principles of psychology,* New York, Henry Holt.

KASTNER, S. & UNGERLEIDER, L. G. 2000. Mechanisms of visual attention in the human cortex. *Annu Rev Neurosci,* 23**,** 315-341.

LABERGE, D. 1983. Spatial extent of attention to letters and words. *Journal of Experimental Psychology: Human Perception and Performance,* 9**,** 371-379.

LEE, J. & MAUNSELL, J. H. R. 2009. A normalization model of attentional modulation of single unit responses. *PLoS ONE,* 4**,** e4651.

MAY, P. J. & TIITINEN, H. 2009. Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology,* (in press).

MCALONAN, K., CAVANAUGH, J. & WURTZ, R. H. 2008. Guarding the gateway to cortex with attention in visual thalamus. *Nature,* 456**,** 391-394.

MEHTA, M. A., SAHAKIAN, B. J., MCKENNA, P. J. & ROBBINS, T. W. 1999. Systemic sulpride in young adult volunteers simulates the profile of cognitive deficits in Parkinson's disease. *Psychopharmacology,* 146**,** 162-174.

MORUZZI, G. & MAGOUN, H. W. 1949. Brain stem reticular formation and activation of the EEG. *Electroencephalography and Clinical Neurophysiology,* 1**,** 455-473.

NAATANEN, R., GAILLARD, A. W. & MANTYSALO, S. 1978. Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol,* 42**,** 313-329.

NAATANEN, R., JACOBSEN, T. & WINKLER, I. 2005. Memory-based or afferent processes in mismatch negativity: a review of the evidence. *Psychophysiology,* 42**,** 25-32.

NÄÄTÄNEN, R. & PICTON, T. 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of component structure. *Psychophysiology,* 24**,** 375-425.

PETKOV, C. I., KANG, X., ALHO, K., BERTRAND, O., YUND, E. W. & WOODS, D. L. 2004. Attentional modulation of human auditory cortex. *Nat Neurosci,* 7**,** 658-663.

POGHOSYAN, V. & IOANNIDES, A. A. 2008. Attention modulates earliest responses in the primary auditory and visual cortices. *Neuron,* 58**,** 802-813.

POSNER, M. I. 1980. Orienting of attention. *Quarterly Journal of Experimental Psychology,* 32**,** 3-25.

RAUSCHECKER, J. P. 1998. Cortical processing of complex sounds. *Curr Opin Neurobiol,* 8**,** 516-521.

RINNE, T., BALK, M. H., KOISTINEN, S., AUTTI, T. & SAMS, M. 2008. Auditory selective attention modulates activation of human inferior colliculus. *Journal of Neurophysiology,* 100**,** 3323-3327.

ROYE, A., SCHRÖGER, E., JACOBSEN, T. & GRUBER, T. 2010. Is my mobile ringing? Evidence for rapid processing of a personally significant sound in humans. *Journal of Neuroscience,* 30**,** 7310-7313.

SAPER, C. B., SCAMMELL, T. E. & LU, J. 2005. Hypothalamic regulation of sleep and circadian rhythms. *Nature,* 437**,** 1257-1263.

SCHALL, J. D. & THOMPSON, K. G. 1999. Neural selection and control of visually guided eye movements. *Annual Review of Neuroscience,* 22**,** 241-259.

SOKOLOV, E. N. 1963. Higher nervous functions: the orienting reflex. *Annual Review of Physiology,* 25**,** 545-580.

STEPHENS, G. J., SILBERT, L. J. & HASSON, U. 2010. Speaker-listener neural coupling underlies successful communication. *Proc Natl Acad Sci USA,* 107**,** 14425-14430.

TITCHENER, E. B. 1908. *Lectures on the elementary psychology of feeling and attention,* New York, MacMillan Company.

TREISMAN, A. M. 1969. Strategies and models of selective attention. *Psychological Review,* 76**,** 282-299.

VALLAR, G. & PERANI, D. 1986. The anatomy of unilateral neglect after right-hemisphere stroke lesions. A clinical/CT-scan correlation study in man. *Neuropsychologia,* 24**,** 609-622.

WINER, J. A. 2006. Decoding the auditory corticofugal systems. *Hearing Research,* 212**,** 1-8.

WOLDORFF, M. G., GALLEN, C. C., HAMPSON, S. A., HILLYARD, S. A., PANTEV, C., SOBEL, D. & BLOOM, F. E. 1993. Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proc Natl Acad Sci USA,* 90**,** 8722-8726.

WOMELSDORF, T., ANTON-ERXLEBEN, K., PIEPER, F. & TREUE, S. 2006. Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. *Nat Neurosci,* 9**,** 1156-1160.

YANTIS, S., SCHWARZBACH, J., SERENCES, J. T., CARLSON, R. L., STEINMETZ, M. A., PEKAR, J. J. & COURTNEY, S. M. 2002. Transient neural activity in human parietal cortex during spatial attention shifts. *Nature Neuroscience,* 5**,** 995-1002.

7 Memory and Learning

Everyone has probably at some wondered how it is possible to recall certain, highly significant, events from one's childhood. Even after tens of years it is possible to remember a wealth of vivid details from such events: sights, smells, sounds, and what took place. And while most of the childhood memories seem to have faded away, it is possible to recall in detail what took place in one's life during adulthood several years before. In fact, it has been often remarked that, unlike the hard disk of a computer, the human memory seems to be of unlimited capacity not to mention lasting much longer than an average computer hard drive. How this is possible, given that the amount of memories that accumulate over the human lifespan is really huge, has been one of the central research topics in human memory research.

Similarly, skills that we learn, such as riding a bicycle, ice skating, or speaking a second language, while eroding if not used in a long time, are quickly revived when one needs them. Another quite amazing commonplace observation of human memory is that, despite the huge amount of information stored in memory, it is still possible to tell right away if one does not know something. For instance, if one is asked for the phone number of some person in another country whom you don't know, it takes only a split second for you to know that you do not have that information in your memory. This ability is called meta-memory.

In fact, those who do not have first-hand experiences of patients with memory deficits might not readily come to appreciate the amazing nature of memory. Imagine what would happen if one would loose memory of one's life events: who he/she are, where he/she lives, how he/she grew up, and memory of his/her friends, and family. One would be utterly stranded under such circumstances and it would take a long while to get back to even close to normal life. On the other hand, imagine a condition in which one would not able to store any new memories, but would constantly wonder where one is and whether one has met before the persons who are present in the same room, even though one might have just a moment before conversed with them.

As will be discussed later in this chapter, these types of memory deficits are called amnesias and they follow from injuries, in some cases very circumscribed ones, to specific structures of the brain. In fact, the human memory is often broken down to three processes: encoding of memories, storage of memories and retrieval of memories (see **Figure 7-1**) and each of these processes can be separately disturbed following brain damage. In the following, however, before describing associations between brain lesions and amnesias, central principles of human memory functions are introduced that have been formulated based on findings from behavioral studies.

Figure 7-1. Memory can be broken down to separate processes of encoding, storage, and retrieval. Each of these processes can be separately damaged following brain injury or disease.

7.1 Memory and learning: behavioral observations

Much like in the other areas of cognitive neuroscience, results from behavioral studies have greatly helped elucidate the nature of memory functions and the underlying neural mechanisms. These studies have revealed the multifaceted nature of memory; indeed there are multitude of memory-related phenomena that range from the forming of associations between rewarding/harmful and neutral stimuli, to memories of variable durations, some lasting for a fraction of a second and some others for a lifetime. Behavioral studies have also answered the puzzling question of how it is possible to store the vast amount of information that accumulates over one's lifetime. In the following, studies that have investigated the forming of simple associations between rewarding stimuli and neutral stimuli will be described first.

Classical conditioning

In the classical studies conducted by Russian scientist Ivan Pavlov, it was observed how a neutral stimulus, when paired with a stimulus that causes the dog to respond, become associated with the response producing stimulus (Pavlov, 1927). After the association has formed, it is then sufficient to present just the neutral stimulus to elicit the response. This phenomenon is called classical conditioning. Specifically, Pavlov observed in his experiments that presenting food to a hungry dog caused the dog to salivate. This is called the unconditioned stimulus. In contrast, ringing a bell, which was the neutral stimulus, did not cause salivation response in the dog. However, when the bell was rung just before presenting the food, Pavlov noticed that the dog started salivating to the ringing of the bell *per se* (see **Figure 7-2**). This observation led to the formation of a scientific discipline called behaviorism, where behavior was thought to be the result of learned associations and reflexes. Behaviorism flourished until the emergence of information processing psychology (which somewhat later was named cognitive psychology) that preceded cognitive neuroscience.

Behaviorism and operant conditioning

The tradition of behaviorism has contributed to cognitive neuroscience in the form of several highly significant principles. Tedious and detailed studies on how different rewards and punishments shape behavioral responses to certain stimuli, called operant conditioning, still today constitutes an important method in animal research that does yield valuable information. Specifically, by giving rewards to behaviors such as depressing a lever that is in the cage where an animal is kept, the animal quickly learns that behavior. The findings from experiments that have used operant conditioning have also motivated relevant studies on human perceptual and cognitive functions. In animal studies, rewards are typically food and sweetened drinks that the animals prefer while for humans obtaining money and social acceptance are strong rewards. The effects of rewards on behavior also explain how drug and alcohol problems ensue, as these substances chemically stimulate the brain's reward circuits.

Extinction of learned behaviors

When the reward is taken out, the animals slowly stop attempting. This is called extinction of behavioral response. Curiously, if the reward is only given intermittently, extinction occurs very slowly (Jenkins and Stanley, 1950). The animals keep on trying for a much longer time than in the case the reward was given to each lever press and then ceased. These observations help explain how addictions develop when, for instance, one is gambling: the reward (winning money) occurs relatively seldom, but the fact that it does occur, results in the gamblers keeping on trying even if they have already lost some money.

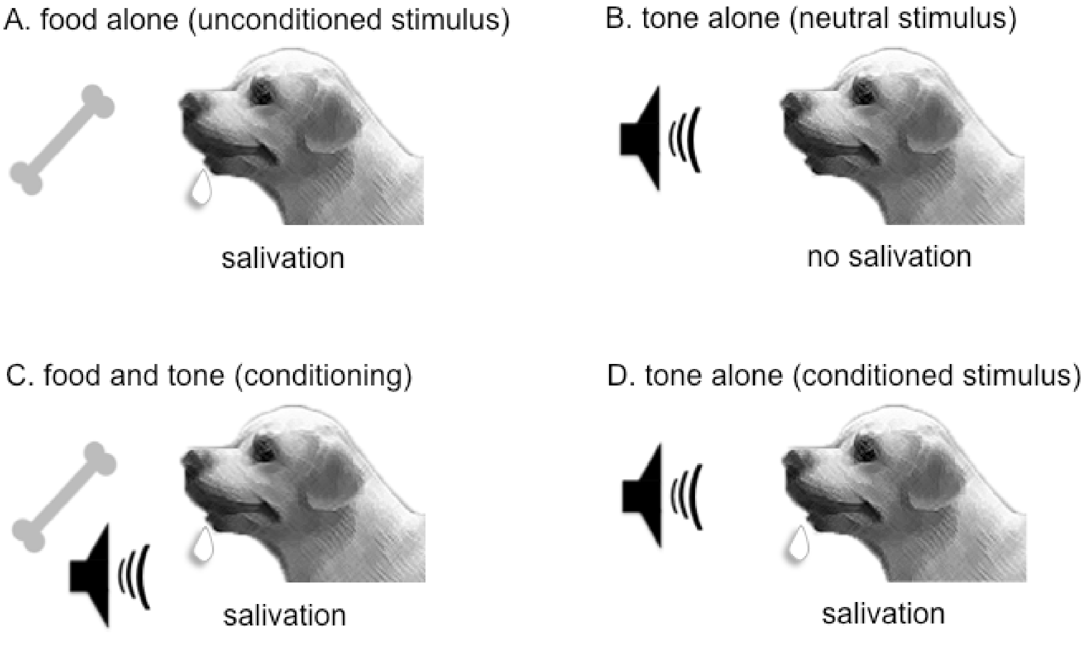


Figure 7-2. Classical conditioning. In early studies conducted by a Russian scientist Ivan Pavlov a dog was A) presented with food (unconditioned stimulus) that caused the dog to salivate. B) The ringing of a bell, when presented without any other stimuli as a neutral stimulus did not cause salivation response in the dog. C) The sound was then presented just prior to the offering of food (*i.e.*, the sound was conditioned with the food stimulus) and, D) when the tone was presented after such paired presentations with food, the sound alone caused the dog to salivate. Thus, the tone had become a conditioned stimulus: a memory trace/representation had formed that associates the sound with food.

Punishment works differently than rewards

Punishments, such as an uncomfortable electric shock, lead to formation of memory traces in an anterior-medial temporal lobe structure called the amygdala. As a result, the animal responds with a so-called fight or flight response to the presentation of the neutral stimulus that has been paired with the shock (the amygdala and conditional learning of emotional responses, as well as development of motivation and addictions are covered in more detail in Chapter 11). In fact, studies on the neural basis of decision-making, motivation, and development of addictions constitute good examples of research questions where the conditioning paradigms are highly useful even today. For a recent review of such studies, see (Balleine et al., 2007).

Three-stage model of human memory

The memory traces that form during classical and operant conditioning constitute long-term, or even permanent/life-long, storage of memories. However, this is not the only type of memory. It was suggested already by William James in the 19th Century that there are two memory systems with different durations and storage capacities (James, 1890). He called the contents of consciousness primary memory, and longer-duration non-conscious and large capacity storage of memorized events and facts secondary memory. Experimental psychology has taken this stance further, by distinguishing between very short-duration sensory memory, limited capacity short-term memory, and long-term memory stores (Atkinson and Shiffrin, 1968). This multi-store model of memory is illustrated in **Figure 7-3** and the central aspects of the model are still used in modern cognitive neuroscience.

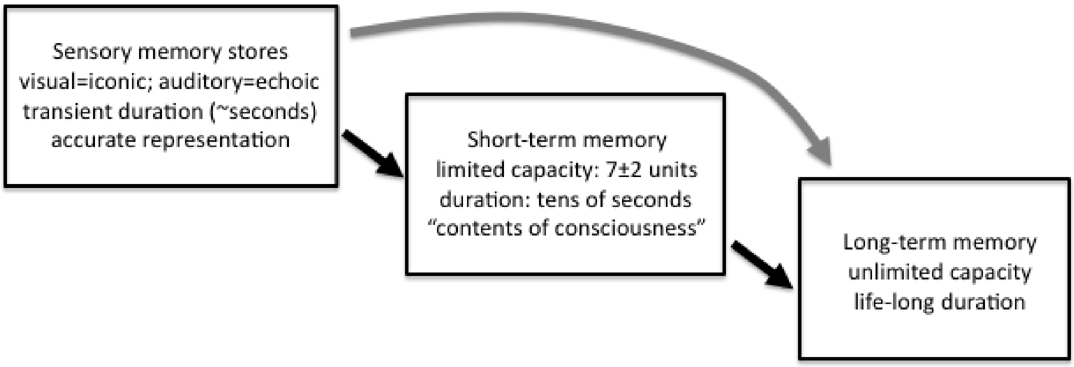


Figure 7-3: The classic three-stage model of human memory. From left to right: incoming stimuli are first coded to sensory memory that are highly accurate representations of the features of sensory input. It has been assumed that there is very little semantic information in sensory memory stores. In the classic memory models, the visual sensory memory called iconic memory is somewhat shorter in duration (~500 ms) than the auditory sensory memory (~seconds) that is called echoic memory. The short-term memory, in turn, is a limited-capacity store where there is no longer an accurate representation of sensory features, but rather memory is represented as semantically meaningful memory units. It is assumed that humans can store 7±2 units at a given time. The short-term memory is assumed to be of duration of several seconds up to tens of seconds, with active rehearsal preventing information from being displaced. The long-term memory has been assumed to be of unlimited capacity and of life long duration. Patient studies do indicate that information can flow directly from sensory memory to long-term memory (after (Atkinson and Shiffrin, 1968)).

Visual and auditory sensory memory stores: iconic and echoic memory

In the three-stage model of human memory, sensory stimuli are first coded into sensory memory that last for a very limited duration. Studies that first suggested the existence of sensory memory stores were conducted in the visual modality (Deutsch and Deutsch, 1963). Sets of digits were flashed for a brief duration (500 ms) followed by a masker stimulus (see **Figure 7-4**). Under such conditions, the subjects were able to recall only a few of the presented digits as the masker stimulus prevented them from utilizing retinal afterimages to aid in recall.

When indicated by an arrow to recollect specific portions of the screen at each presentation the subjects were able to, in parts, recall the whole array of digits. This was taken as evidence suggesting that the sensory memory holds the whole array of digits but for such a limited duration that subjects do not have time to recall everything. In general, it is assumed that the visual sensory memory, also often referred to as the iconic memory, holds an accurate and very short-lived representation of physical visual stimulus features (*i.e.*, it is assumed that sensory memory does not hold semantic information).

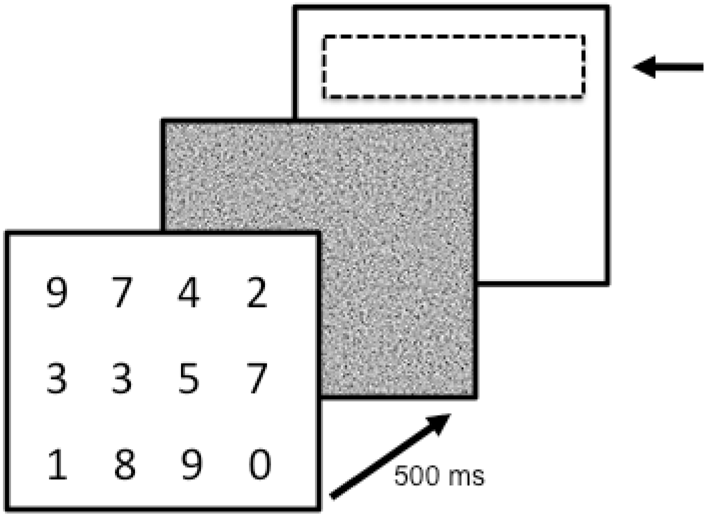


Figure 7-4. Shown is the experimental setup that was utilized in studies of the visual sensory (or iconic) memory. A set of digits was presented first for a very brief duration, followed by a masker stimulus and request to recall immediately the digits that had been shown. The masker stimulus was used to prevent the subjects from using retinal afterimage to recall the digits. When the subjects were requested to recall all the digits at once, they could only remember a small portion of the digits shown. However, when indicated by a cue arrow to recollect specific portions of the screen at each presentation the subjects were able to, in parts, recall the whole array of digits. This indicates that the iconic memory does hold the whole visual field of view, but for such a brief duration that subjects are not able to recall everything before the sensory memory representation decays or is displaced.

In the auditory modality, studies have been conducted where auditory stimuli are presented in the background while the subjects are conducting another task such as reading a book. Intermittently, the subjects are then prompted to recall the most recently presented auditory stimulus (Sams et al., 1993). From such studies it has been concluded that the auditory sensory memory duration is somewhat longer than that of the iconic memory, lasting for several seconds. However, if masker stimuli are used analogously to the studies on iconic memory, the estimated auditory sensory memory duration is much shorter.

Based on behavioral studies it has also been suggested that the sensory memory systems consists of very short-lived (on the order of 500 ms) highly accurate representations of the physical stimulus features in both the auditory and visual modalities, and of a longer-duration (from seconds to tens of seconds) sensory memory that would already retain more abstract representations of the stimuli (Cowan, 1995). As a concrete example of how echoic memory is utilized in everyday life, amidst reading a newspaper one is able to recall the last couple of words of what was being said after shifting of attentional focus to the verbal demands of one's spouse.

Short-term memory

The stage of memory encoding that is subsequent to sensory memory is called short-term memory. The short-term memory also differs from sensory memory stores in that its capacity is limited to 7±2 "items". The memory representations (or items) are also semantic in nature, in contrast to the physical features assumedly stored by sensory memory. The memory items of short-term memory depend on semantic knowledge of the individual. For a person who has never heard of personal computers before, the letters "PC" are two items, whereas for the more computer-able individuals, these letters comprise a single short-term memory item. This, in fact, is one of the reasons why it is easier for experts to handle large amounts of data – to them the short-term memory items are more extensive constructs wherein the lay persons have to handle the same information with much larger number of memory items.

Short-term memory representations are also longer-lived than the sensory memory representations. They last for several seconds or tens of seconds, and if the subjects rehearse the memory items, they can be kept in short-term memory much longer. The duration of the short-term memory store has been investigated behaviorally using, for example, a task called Brown and Petersen auditory consonant trigrams procedure (Peterson and Peterson, 1959, Lezak et al., 2004). In this task, the memory duration is investigated by presenting to subjects non-meaningful consonant trigrams (*e.g.*, "SWK"), followed by a random number from which the task of the subjects is to count down at steps of three (*i.e.*, if 379 is given, the subjects are to count "376, 373, 370…"). The purpose of the backwards-counting task is to prevent subjects from rehearsing the three to-be-memorized consonants. Then the experimenter will, either after 3, 6, 9, 12, 15 or 18 seconds has passed, prompt the subject to recall the consonant trigram.

By presenting a sufficient number of trials, an average forgetting rate can then be calculated as a function of time to recall. It has been shown that while more than half of the items are remembered after three seconds, the percentage of trigrams recalled is on the order of 15% after delay of 18 seconds. Another source of information on the rate of forgetting comes from studies where the task of the subjects has been to recall lists of various items presented to them. With longer lists, it has been noted that the first one or two items on the list, and the last few items, are recalled most accurately. It has been speculated that this is because the first couple of items have been recited and transferred to long-term memory, whilst the items towards the end of the list can be recalled from the short-term memory.

Furthermore, in a so-called Sternberg task, it has been studied how items are recalled from short-term memory (Sternberg, 1966). In this task, experimental subjects are first presented a string of letters, each comprising a memory item. Then, a target letter is presented, with the question of whether it was among the list of to-be-remembered items. The results of these experiments have suggested that search from short-term memory occurs in serial-exhaustive manner. That is, the speed of responding did not depend on the serial position of the target item on the list, but only on the length of the string of letters that was memorized. Thus, the search is not terminated upon finding the target letter. In this task, the speed of memory search appeared to be 25-30 symbols per second.

These studies are good examples of how behavioral and reaction time measures have been utilized to learn about memory functions. It is, in fact, rather amazing how much have been learned about the human cognitive functions and the underlying neural mechanisms by drawing inferences based on behavioral measures without modern neuroimaging methods that allow one to measure brain activity. Since the early days of information-processing psychology, the memory models have been significantly developed from the three-store model. As one significant subsequent advance in the field, Alan Baddeley introduced the concept of working memory to account better than the classical short-term memory model for the type of flexible processing of transient memory representations that takes place under ecologically valid conditions (Baddeley and Hitch, 1974, Baddeley, 1992). This multiple-component working memory model is described next.

The multiple-component model of working memory

The most significant difference between short-term memory and working memory concepts is that the working memory model assumes active manipulation of both sensory information and information that is searched from long-term memory. The multiple-component working memory model consists of a central executive that utilizes two subsystems called the visuo-spatial sketchpad and phonological loop (see **Figure 7-5**). In the model, the central executive is thought to be responsive for filtering information from both the senses and from long-term memory to (and actively manipulating information in) these two subsystems.

Solving arithmetic problems is a good example of the need for actively manipulating information in the flow of consciousness and also of the difficulties posed by the inherently limited capacity of the working memory. Another commonplace example is rehearsing (or looping) a phone number in working memory, until one gets to dial the number. In the model, the phonological loop is handling verbal (such as words, sentences, verbal concepts), and the visuo-spatial sketchpad non-verbal information, such as routes when driving around in a city and visual imagery. In 2000 Baddeley expanded his model by adding a third subsystem called episodic buffer (Baddeley, 2000).

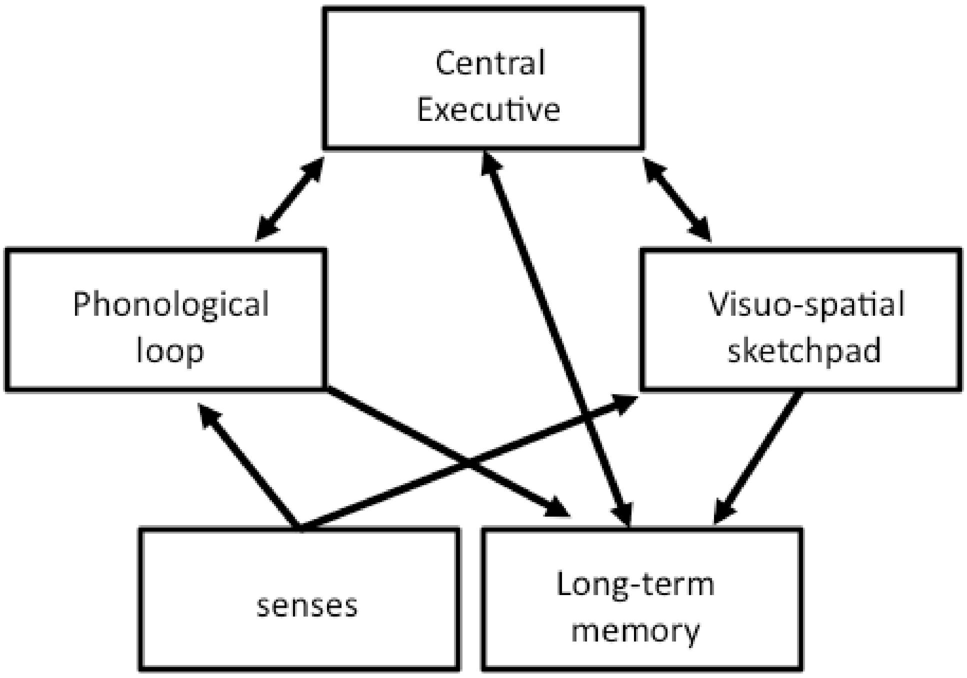


Figure 7-5. Schematic illustration of the central tenets of the working memory model. The central executive is using two short-term memory subsystems, the phonological loop (for verbal/language information) and visuo-spatial sketchpad (visual and spatial information) to actively manipulate information that is needed in concurrent tasks. In the model, the central executive also retrieves information from long-term memory and filters task-relevant sensory information.

The long-term memory

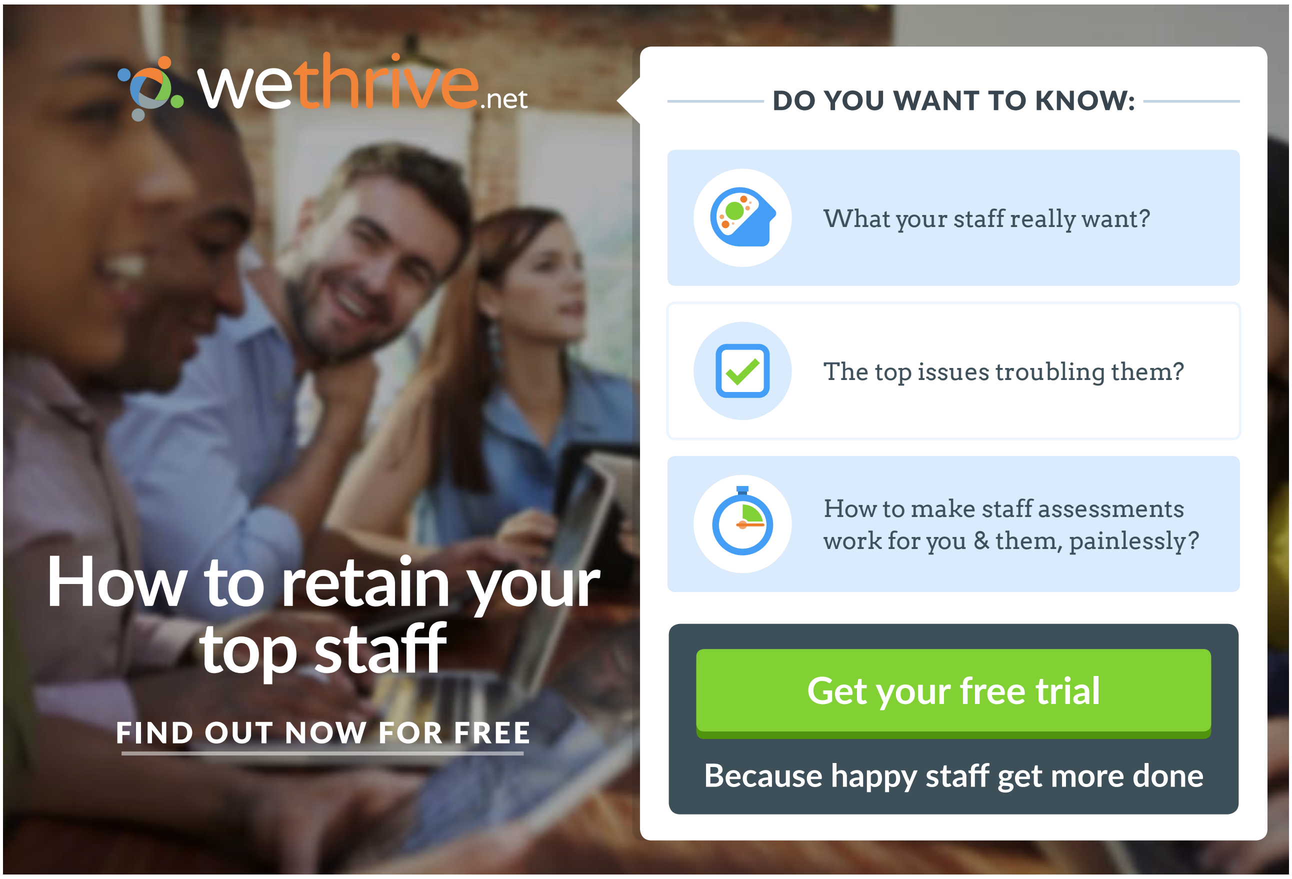
Long-term memory is responsible for up to lifelong storage of massive amounts of information. In long-term memory terminology, episodic memory (*i.e.*, memory of life-events that occurred to oneself) and semantic memory (common knowledge of the world) is jointly called declarative memory. Non-declarative memory, in turn, refers to both procedural memory (*i.e.*, learning of motor skills such as riding a bicycle) and sensory learning (*e.g.*, learning to discriminate between two musical instruments). The latter has been shown to be associated with receptive field changes in sensory cortices that allows discrimination of specific sounds (for a recent review see (Jaaskelainen et al., 2007)).

As will be described in more detail below, the hippocampus is a central anatomical structure for consolidation of episodic and semantic information to long-term memory storage, and while the neural mechanisms of long-term storage are not precisely known, it is widely assumed that alterations in synaptic strengths in the brain, by changing the way that signaling flows in networks of neurons, underlie long-term storage. But how is it possible that we have a vast, life-long storage of all the information that we have encountered or learned about during our lives? The key to understanding this is that information is both stored and retrieved with the help of memory representations called schemas.

Schemas explain long-term memory capacity

There are a few findings that help one to understand the role played by schemas in long-term memory. In one study, experimental subjects entered a graduate student's office or a preschool classroom. Afterwards they were requested to recall the items that were present in the room. The recall was characterized by two principles. First of all, items that are typical to each setting, such as phones in the office, are remembered even if the office did not furnish any. Secondly, items that are highly untypical but perceptually salient, such as a skull placed on shelve or table in the office, are remembered very well. This suggests that only certain significant key elements of each life event are stored in long-term memory. Recall of a given life event is then a reconstruct based on common knowledge and the few significant details, rather than an accurate recollection of each detail (Pezdek et al., 1989).

Importantly, these findings reflect the reconstructive nature of memory, a principle that was introduced by Frederick Bartlett (Bartlett, 1932). Bartlett carried out studies in which subjects read a Native American story about ghosts and the retold the story. Due to variability in the cultural backgrounds of the subjects, they changed, added, and omitted details of the story, as they could not understand specific aspects of the story. Another example of the reconstructive nature of memory is provided by studies where subjects have been asked to reproduce a drawing with a delay. The reproduced drawing is then shown to another subject who memorizes it and draws it after a delay. When this is repeated across several subjects, the reconstructive nature of memory is readily revealed, as shown in **Figure 7-6** below; it can be seen how the original drawing is modified based on schema-guided recall of the drawing when the subjects are reproducing the figure.



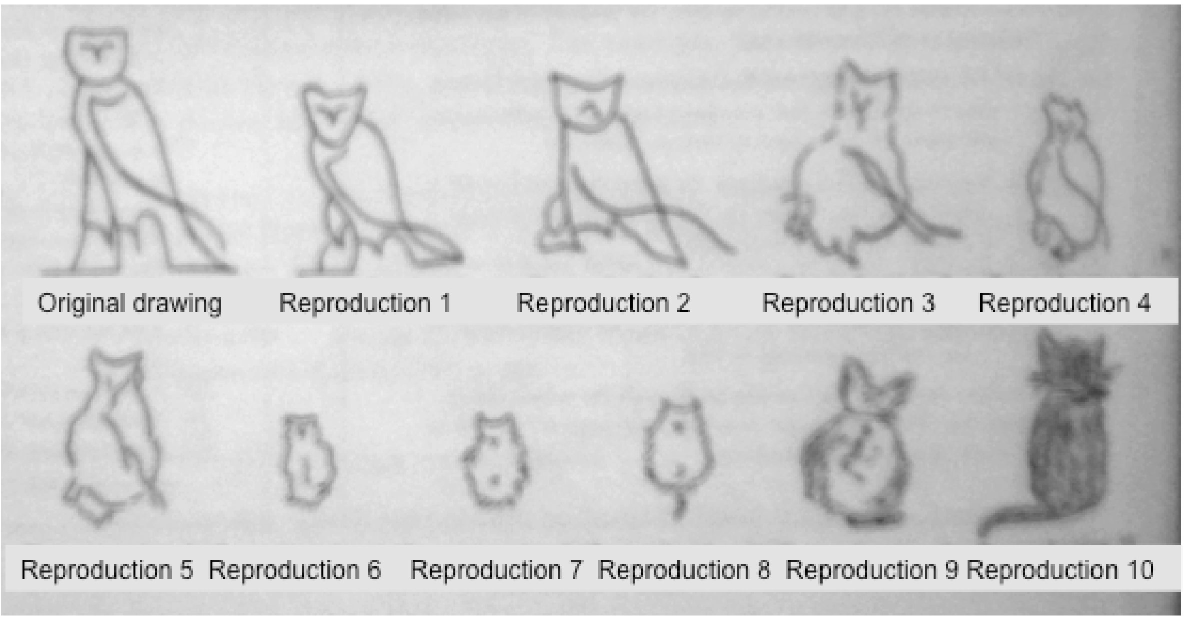


Figure 7-6. Reconstructive nature of memory as disclosed by serial picture reproduction experiments. Here, a subject is first asked to reproduce a drawing from memory, which is then presented to another subject who in turn reproduces from memory the picture he/she has seen, which is then seen and reproduced by a third subjects, and so on, up to 10th subject. It can be seen how the drawing is altered, presumably guided by schema-based recall of the drawing by each of the subjects. This has been taken as evidence supporting the reconstructive nature of memory (*i.e.*, that all the details are not stored but rather a few significant key features that are then retrieved with the help of schema-based general knowledge that fills in the missing details) (adapted from (Bartlett, 1932)).

Semantic-level processing facilitates memory encoding

Factors that predict successful memory encoding have been investigated using tasks where subjects are processing verbal material using various strategies. As one example of the effect of use of memorization strategies during encoding, when healthy volunteers were presented a list of words and they were to either 1) tell whether the words were written with capital letters or not, 2) tell whether the words rhymed together, or 3) evaluate the words semantically (*e.g.*, "does it fly"). In these types of levels-of-processing experiments, the best recall of the list of words occurs when the words have been processed semantically, and worst when the processing has been very superficial, such as when telling only whether the word is capitalized or not (Craik and Tulving, 1975). In subsequent neuroimaging studies, prefrontal cortical activations were observed during deep processing and enhanced recall of items (Kapur et al., 1994). In addition to semantic level of processing, emotions also enhance memory encoding, but primarily of details that are associated with the center of the emotional event (Burke et al., 1992). It is possible that dense interconnectivity between amygdala and hippocampus underlie this effect.

Hypnosis and the capacity of long-term memory

It has been sometimes claimed that people would be able to produce highly detailed recollections under the state of hypnosis and, thus, that memory encoding would be highly efficient and failing to remember would be mostly due to problems in retrieving memories. However, in a critical study, subjects were first shown a short movie of a car accident. They were then asked, when under hypnosis *vs*. in a normal state, what the register plate number of the car was. Under the normal state, subjects said that they are unable to recall the plate number, but when under hypnosis the subjects were recalling the register plate number. However, the register plate number was never shown in the movie (Putnam, 1979). This example demonstrates how the subjects, when under hypnosis, enter a state of mind where their ability to judge whether the recollections are correct is reduced, thus likening their recollections to confabulations that patients with prefrontal cortical damage produce (see below). These findings also support the view that long-term memory does not store detailed accounts of life events but rather few significant details that are then utilized, together with schema-based common knowledge of such events, to reconstruct the memories.

Is forgetting due to decay or interference?

The failure of memory, forgetting, is another topic that has inspired considerable amount of research. One of the earliest behavioral studies of human memory and rate of forgetting were conducted by Herman Ebbinghaus in 1885 (see (Ebbinghaus, 1964)). He used himself as an experimental subject, devised lists of pseudowords consisting of strings of random letters that carried no semantic meaning, and then tested his ability to remember the lists with varying delays to recall. Ebbinghaus observed that the rate of forgetting is quite steep, that memorization is better after an overnight sleep, and that the amount of lists that have to be learned have a significant effect on the rate of forgetting.

Indeed, one of the central questions concerning forgetting from memory has been whether information is lost from long-term memory due to decay of memory traces, or due to interference. The latter means that one cannot find the relevant information due to there being too much similar overlapping information. For instance, it is possible to remember even after a delay of years what one had for breakfast in some hotel that one has only visited once before, whereas telling what one had for lunch three months aback in the local cafeteria one eats in daily would be impossible.

Another indication that forgetting might be mainly due to interference comes from studies where the effects of contextual cues have been shown to aid recall. For instance, divers can better recall items that they learned while under water when they are diving, as opposed to when they are on the surface (Godden and Baddeley, 1975). Similarly, it is possible that the rapidly changing contexts that occur when one grows up (everything becomes smaller in size relative to oneself, one's thought processes develop giving a different kind of perspective into things) explains why so relatively few details can be remembered from childhood.

7.2 The neural basis of memory

Thus far in this chapter, the memory functions have been described without much reference to the underlying neural mechanisms. Thanks to a body of literature on specific memory deficits in neurological patients with circumscribed brain lesions, there is today highly specific information on the neural basis of memory. Further insights into the neural basis of memory have been recently provided by neuroimaging studies. These findings will be described in the following, beginning with description of memory deficits, amnesias, that have been observed in brain damaged patients.

Retrograde amnesia

The inability to recall events that happened prior to a brain injury or concussion is called retrograde amnesia. Typically retrograde amnesia occurs following a head trauma with a loss of consciousness and it can initially extend weeks, months, or even years back in time from the time of the accident (Fisher, 1966). However, this often concerns only memory of events that have occurred to oneself, which is referred to as episodic memory. Knowledge of things, such as recognition of scissors and what one can do with them, which is referred to as semantic memory, is often not affected in retrograde amnesia. Further, procedural memory (*i.e.*, skills such as the ability to ride a bike, or knit) is not lost in retrograde amnesia. These observations have led to episodic, semantic, and procedural memories being viewed as distinct types of memory with separate underlying neural mechanisms.

With gradual recovery from retrograde amnesia, memories that are furthest away from the time of the accident are retrieved first. It is typical, however, that events taking place just prior to the accident are never retrieved. From this observation it has been inferred that it takes a while, typically from tens of minutes up to a few hours, for memories to consolidate, and those memories that have not consolidated before the trauma occurs can be permanently lost. In another line of research, electrical stimulation of a specific brain structure called hippocampus has been observed to interfere with memory consolidation, producing retrograde amnesia (Sideroff et al., 1974). However, there is evidence indicating that memories are not stored in hippocampus, since damage to hippocampus does not necessarily result in loss of acquired memories, but rather the hippocampus seems to highly important for memory encoding (see below).

Memories are stored as distributed representations

It is not precisely known where memories are stored in the brain, however, it is a widely held view that memories are coded as changes in synaptic strengths that are distributed across the brain. As an analogy, this has sometimes been compared to depressions on a side of a hill that streams of water make when it rains. Then, when it rains again, water runs through the same depressions. It makes a lot of sense to store information in a distributed fashion, rather than assigning each memory a unique unit in the brain. This is for the simple reason that by using distributed code it is possible to represent more items.

The power of distributed representations is schematically illustrated in **Figure 7-7**. Consider that there is an array of 16 memory units that are in the figure represented as circles. If each of them was assigned a unique memory, one could only represent 16 memories. However, if one uses distributed code, where the pattern of "on" and "off" elements is what codes the memory as shown in the figure, one can represent up to 65536 items. It is currently believed that this does not only apply to long-term memory, but rather that distributed representations might also underlie representations of perceptual objects in higher-order visual cortices (Ishai et al., 1999, Cox and Savoy, 2003) and also the short-term memory stores that are described below have been speculated to use distributed representations (Jaaskelainen et al., 2011).

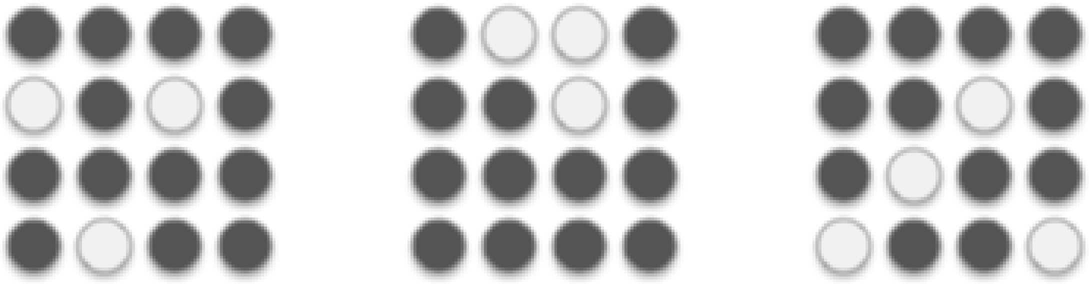


Figure 7-7: An example of distributed representations. Here, three arrays of 16 elements are shown with three different distributed representations. Each circle is a unit in the array. The light circles represent "on" units and the dark ones "off" units. If one was to use these units in such a fashion that the "on" state of each of the units would represent a single memory item, only 16 memories could be stored. However, when memory representations are formed by using the units of the array in specific combinations of "on" and "off" states, as exemplified in the figure with three such distributed representations, more than 64000 different representations are made possible with these 16 units. Given that the human brain has approximately 10^{11} neurons, the capacity to represent memories using a distributed code is in principle very large.

Hippocampus damage results in anterograde amnesia

While hippocampal damage (often) does not result in retrograde amnesia, another type of memory loss results from damage to this medial temporal lobe structure. The now famous patient H.M. had his hippocampus removed bilaterally as a surgical treatment of epilepsy. **Figure 7-8** illustrates the hippocampal lesions of H.M. based on MRI scans taken several years following the neurosurgery. While he was able to remember things that had occurred several years prior to the lesion (events that had taken place in the most recent three years were lost from his memory) H.M. was unable to form any new memories (Scoville and Milner, 1957).

As an example of the memory-forming problems exhibited by patient H.M., mere minutes after talking with a person H.M. could not remember having ever met that person before. He was also unable to carry any conversations since he quickly forgot what had been said previously. He described feeling confused "like waking from a dream… every day is alone in itself …" The memory deficits of patient H.M. have been investigated thoroughly and for an extensive recent review on the findings on patient H.M., see (Squire, 2009). This type of memory loss, where events prior to accident are recollected but events after the accident cannot be memorized, is called anterograde amnesia.

Curiously, while the ability of H.M. to encode episodic memories was lost, his procedural memory seemed to be relatively unaffected. When H.M. was learning a new sensory-motor skill, copying a drawing when he could only see his hand and the drawing through a mirror, he exhibited a relatively normal learning curve (Milner, 1962). However, characteristic of his memory problem, during each of the practice sessions, he was unable to remember that he had ever carried out the task before. This is another indication that learning of skills and encoding of episodic memories are distinct neural processes.

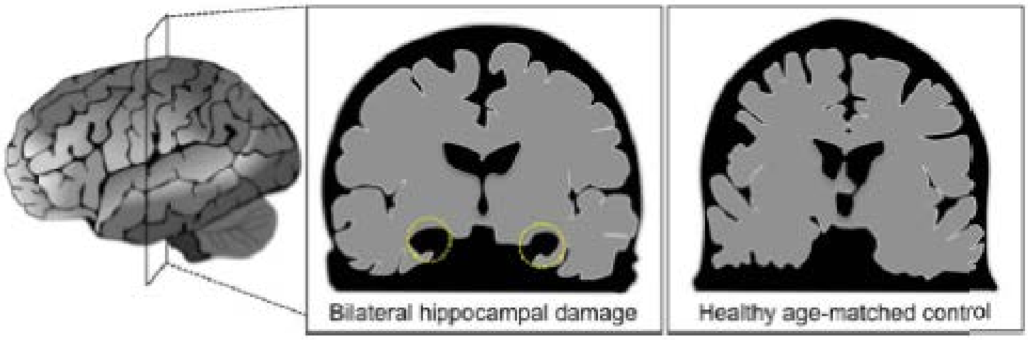


Figure 7-8: An illustration of the bilateral hippocampus damage in patient H.M. that led to anterograde amnesia, drawn based on MRI scans reported in (Corkin et al., 1997). On the left is shown the approximate location of the coronal plane on which the damaged hippocampi are illustrated in the middle. The yellow circles indicate the damaged hippocampi. Coronal plane showing the brain of a healthy age-matched control is displayed on the right.

Lesion of specific hippocampus regions results in anterograde amnesia

While in the patient H.M. the hippocampi were destroyed extensively in both the left and right hemispheres, similar anterograde amnesia has been observed in another patient with a lesion of specific regions of hippocampi termed CA1 and CA2 (Kartsounis et al., 1995). This tentatively suggests that within the hippocampus there are regions that are especially necessary for encoding/consolidation of episodic memories. Furthermore, there is evidence suggesting that when only one of the hippocampi is damaged (*i.e.*, unilateral rather than bilateral damage) the types of memory deficits that ensue depend on the side of the lesion (Cabeza and Nyberg, 1997).

Left hippocampus lesions seem to result in memory deficits involving verbal and language materials such as memorizing stories that one hears. Right hippocampus lesions, in turn, result in memory deficits involving spatial material such as deficits in ability to remember how to find one's way in an unfamiliar environment such as an office building or town that one has not been into before. Thus, while bilateral hippocampus damage results in anterograde amnesia, unilateral damage results in reduced ability to consolidate more specific types of memories.

Hippocampal long-term potentiation

At the neurophysiological level, there is a phenomenon called long-term potentiation that was first found in the hippocampus, but has been subsequently described also in other structures such as the amygdala (see Chapter 11). The hippocampal long-term potentiation takes place when there is synchronous depolarization of pre- and post-synaptic neurons *via* glumatergic activation of N-methyl-D-aspartate receptors that results in chemical changes that strengthens the synaptic connections between the neurons. The long-term potentiation has been studied extensively as a potential neural correlate of memory formation (Bliss and Collingridge, 1993), and is considered to underlie synaptic plasticity more generally (for an introduction to mechanisms of brain plasticity, see Chapter 8).

Acetylcholine is a central neurotransmitter for memory consolidation

Another way that anterograde amnesia can be induced is pharmacologically by drugs that block muscarinic acetylcholine receptors. The acetylcholinergic tracts in the brain are depicted in **Figure 7-9**. Analogously, patients with Alzheimer's disease suffer from anterograde amnesia and lesions of acetylcholinergic tracts from nucleus basalis of Meynert have been implicated as a potential underlying mechanism. Indeed, drugs that enhance the amount of acetylcholine in synapses, acetylcholinesterase inhibitors, alleviate the anterograde memory deficits in these patients (for a recent review, see (Birks, 2009)). The role of acetylcholine is, however, not limited to consolidation of episodic memories, but animal studies suggest that also sensory learning (learning to discriminate between two sensory stimuli such as timbres of musical instruments) is dependent on the ascending cholinergic inputs (Froemke et al., 2007).

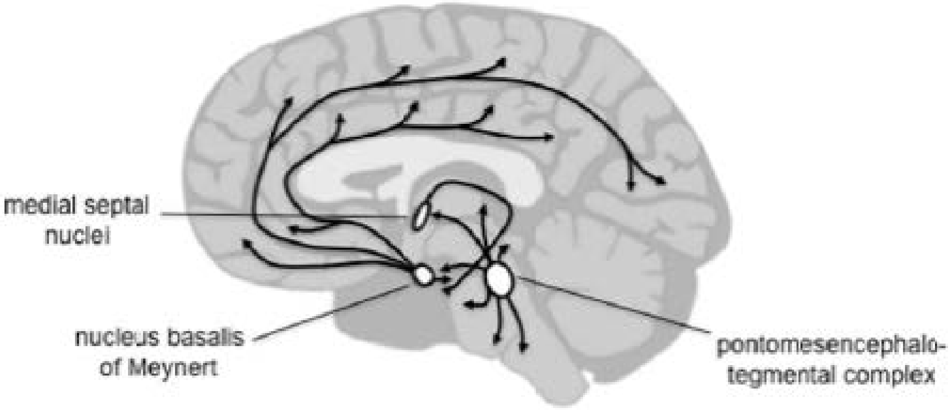


Figure 7-9. Shown are the cholinergic systems of the human brain. The nucleus basalis of Meynert, as well as medial septal nuclei, project widely to the cerebral cortex, including the hippocampus. The pontomesencephalotegmental complex projects to the thalamus and the brain stem.

Problems of encoding and retrieval after frontal lobe damage

Amnesic conditions where there are problems retrieving memories also often occur following damage to frontal lobes of the brain, especially to the prefrontal cortical areas. Damage to these areas does not seem to disturb the memory consolidation *per se*, but reduced ability to use mnemonics (*i.e.*, memory cues and memorizing strategies to aid learning) during both encoding and recall seems to contribute to memory difficulties in these patients (Gershberg and Shimamura, 1995). The way that this has been studied is comparison between to what extent the patients are able to recognize stimuli that have been presented to them before *vs*. to what extent they can remember those stimuli without seeing them—remember that patient H.M. was not able to even recognize stimuli that he had seen before.

Further, what is also striking in some prefrontal cortical damaged patients is the inability to evaluate which memories are real ones and which ones are products of their imagination that have never really occurred (for a recent review, see (Schnider, 2003)). This is called confabulation and it was first described in alcoholic patients (Korsakoff, 1889), in whom prolonged drinking together with thiamine deficiency leads to frontal lobe damage. Confabulation is manifested by patients telling lively about events that have occurred to them that turn out to be false memories when critically examined.

In neuropsychological testing memory can be assessed by telling patients short stories that they are requested to recall both immediately and with a delay of 20-30 minutes without prior warning. Memory consolidation is then assessed as the difference between immediate and delayed recall, specifically how many critical elements of the story are recollected. Confabulation is in such tests manifested by intrusion of new elements into the story and gross alteration of the plot of the story, without the patient being aware of the false nature of the recollections.

Other memory deficits associated with the frontal lobes

In addition to failure to use memory cues and confabulation, prefrontal cortical damage is also associated with other types of memory deficits. These patients often have problems in judging the temporal order of memories (Shimamura et al., 1990). They might recall that a certain event has occurred, but are unable to tell correctly when it has occurred. Overall, they have difficulties in placing the memories in the right context, for instance, remembering where they have met a person before. Yet another memory function in which prefrontal cortical areas have been found to play a central role is prospective memory (*i.e.*, "remembering to remember" to do something) (Burgess et al., 2001).

One quite interesting finding is deficiency in release from proactive interference (or inhibition) in frontal-lobe damaged patients (Thompson-Schill et al., 2002). When a healthy person has to learn memory items of a given category, such as grocery store lists, at some point items from preceding lists start to interfere the recall. Upon shifting to another category of memorized items, such as hardware store lists, the grocery store list items cease to interfere. This phenomenon is called release from proactive inhibition/interference (Wickens et al., 1963). In frontal lobe damaged patients, this does not occur as the grocery store list items keep on intruding into the hardware store lists. This suggests that what the frontal lobe patients have difficulty in is the use of hierarchically higher order memory representations or schemas that help organize memory.

Stimulus-specific adaptation supports sensory memory

The primary sensory cortical areas are naturally vital for sensory memory, as lesions of sensory cortex results in loosing the ability to consciously perceive stimuli of the respective sense modality. Neuroimaging studies have provided evidence for the neural basis of sensory memory within the sensory cortices; neurons that respond to specific stimulus features, such as sound frequency, intensity, or a specific phonetic category, are transiently inhibited after the occurrence of that feature. This phenomenon is called stimulus-specific adaptation and the candidate neurophysiological mechanisms include synaptic depression and lateral/surround inhibition (Jaaskelainen et al., 2011).

In animal models, it has been shown that there are continuums of adaptation time-scales in the auditory cortex, ranging from tens of milliseconds to tens of seconds (Ulanovsky et al., 2004). Furthermore, it has been shown that the rate of recovery from adaptation correlates with the duration of auditory sensory memory in humans (Lu et al., 1992). Stimulus-specific adaptation of different parts of the auditory cortex has also been suggested to underlie an event-related potential correlate of sensory memory called mismatch negativity (Jaaskelainen et al., 2004). Taken together, these different lines of research suggest that neuronal adaptation by sensory stimuli in sensory cortical areas could underlie sensory memory, although this is an issue that is currently being debated (Naatanen et al., 2005, May and Tiitinen, 2009).

Neural mechanisms underlying working memory

The neural mechanisms underlying short-term memory or, in more modern terms working memory, have been also studied extensively. It has been shown that there are neurons in the prefrontal cortex of non-human primates that fire continuously during the delay period in a matching-to-sample task (Fuster and Alexander, 1971). This delay period firing has been thought to underlie maintenance in working memory of the target object representation. Human neuroimaging studies have documented the role of prefrontal cortex in supporting the “central executive” component of working memory, for example, orbitofrontal cortex activity has been observed to be associated with selection of relevant information to, and blocking out of irrelevant information from, working memory (Schnider et al., 2000). In brain damaged patients, lesions to left supramarginal gyrus and premotor cortex have been reported to result in deficits in the phonological loop, whereas lesions to posterior parietal areas on the right side of the brain have been reported to result in visuo-spatial sketchpad deficits (Lezak et al., 2004).

7.3 Concluding remarks

Human memory is a multifaceted phenomenon, with distinct neural mechanisms supporting transient sensory memories, manipulation of memories in one’s awareness, and large-capacity long-term storage of memories. It seems that multiple brain areas are involved, with sensory cortices important for sensory memories, prefrontal cortex important for the central executive of working memory, including retrieval of relevant information from long-term memory and prospective memory, and medial temporal lobe structures for encoding/consolidation of memories. Notably, it can be argued that being able to store memories and learn new abilities primarily serves the purposes of survival and prediction of the future; based on what one knows of past events, and being able to utilize that information in predicting what will happen next, helps one to avoid / get prepared for dangerous situations or, in the modern world, enables one to smoothly interact with other members of the society and, overall, manage in daily life. The helplessness of patients who have lost their memory abilities highlights the vital importance of memory functions. One very interesting and important research area in cognitive neuroscience that is intimately linked with the neural mechanisms supporting memory and learning is cortical plasticity; how the developing brain is shaped and how the brain adapts after damage, radical changes in the environment, or a loss of sensory input. Plasticity will be described in the next Chapter of this book.

7.4 References

ATKINSON, R. C. & SHIFFRIN, R. M. 1968. Human memory: a proposed system and its control processes. *In:* SPENCE, K. W. & SPENCE, J. T. (eds.) *Psychology of Learning and Motivation: Advances in Research and Theory.* New York: Academic Press.

BADDELEY, A. 1992. Working memory. *Science,* 255**,** 556-559.

BADDELEY, A. 2000. The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences,* 4**,** 417-423.

BADDELEY, A. & HITCH, G. 1974. Working memory. *In:* BOWER, G. H. (ed.) *The psychology of learning and motivation: advances in research and theory.* New York: Academic Press.

BALLEINE, B. W., DELGADO, M. R. & HIKOSAKA, O. 2007. The role of the dorsal striatum in reward and decision-making. *Journal of Neuroscience,* 27**,** 8161-8165.

BARTLETT, F. C. 1932. *Remembering: a study in experimental and social psychology,* Cambridge, Cambridge University Press.

BIRKS, J. 2009. *Cholinesterase inhibitors for Alzheimer's disease*, The Cohcrane Library.

BLISS, T. V. P. & COLLINGRIDGE, G. L. 1993. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature,* 361**,** 31-39.

BURGESS, P. W., QUAYLE, A. & FRITH, C. D. 2001. Brain regions involved in prospective memory as determined by positron emission tomography. *Neuropsychologia,* 39**,** 545-555.

BURKE, A., HEUER, F. & REISBERG, D. 1992. Remembering emotional events. *Memory & Cognition,* 20**,** 277-290.

CABEZA, R. & NYBERG, L. 1997. Imaging cognition: an empirical review of PET studies with normal subjects. *Journal of Cognitive Neuroscience,* 9**,** 1-26.

CORKIN, S. C., AMARAL, D. G., GONZÁLEZ, R. G., JOHNSON, K. A. & HYMAN, B. T. 1997. H.M.'s medial temporal lobe lesions: findings from magnetic resonance imaging. *Journal of Neuroscience,* 17**,** 3964-3979.

COWAN, N. 1995. Sensory memory and its role in information processing. *In:* KARMOS, G., MOLNAR, M., CSEPE, V., CZIGLER, I. & DESMEDT, J. E. (eds.) *Perpectives of Event-Related Potentials Research (EEG Suppl 44).* Elsevier Science B.V.

COX, D. D. & SAVOY, R. L. 2003. Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage,* 19**,** 261-270.

CRAIK, F. I. M. & TULVING, E. 1975. Depth of processing and the retention of words in episodic memory. *J Exp Psychol Gen,* 104**,** 168-294.

DEUTSCH, J. A. & DEUTSCH, D. 1963. Some theoretical considerations. *Psychol Rev,* 70**,** 80-90.

EBBINGHAUS, H. 1964. *Memory,* New York, Dover.

FISHER, C. M. 1966. Concussion amnesia. *Neurology,* 16**,** 826-830.

FROEMKE, R. C., MERZENICH, M. M. & SCHREINER, C. E. 2007. A synaptic memory trace for cortical receptive field plasticity. *Nature,* 450**,** 425-429.

FUSTER, J. M. & ALEXANDER, G. E. 1971. Neuron activity related to short-term memory. *Science,* 173**,** 652-654.

GERSHBERG, F. B. & SHIMAMURA, A. P. 1995. Impaired use of organizational strategies in free recall following frontal lobe damage. *Neuropsychologia,* 33**,** 1305-1333.

GODDEN, D. R. & BADDELEY, A. D. 1975. Context-dependent memory in two natural environments: on land and underwater. *British Journal of Psychology,* 66**,** 325-331.

ISHAI, A., UNGERLEIDER, L. G., MARTIN, A., SCHOUTEN, J. L. & HAXBY, J. V. 1999. Distributed representation of objects in the human ventral visual pathways. *Proc Natl Acad Sci USA,* 96**,** 9379-9384.

JAASKELAINEN, I. P., AHVENINEN, J., ANDERMANN, M. L., BELLIVEAU, J. W., RAIJ, T. & SAMS, M. 2011. Short-term plasticity as a neural mechanism supporting memory and attentional functions. *Brain Research,* 1422**,** 66-81.

JAASKELAINEN, I. P., AHVENINEN, J., BELLIVEAU, J. W., RAIJ, T. & SAMS, M. 2007. Short-term plasticity in auditory cognition. *Trends in Neurosciences,* 30**,** 653-661.

JAASKELAINEN, I. P., AHVENINEN, J., BONMASSAR, G., DALE, A. M., ILMONIEMI, R. J., LEVANEN, S., LIN, F.-H., MAY, P., MELCHER, J., STUFFLEBEAM, S., TIITINEN, H. & BELLIVEAU, J. W. 2004. Human posterior auditory cortex gates novel sounds to consciousness. *Proc Natl Acad Sci USA,* 101**,** 6809-6814.

JAMES, W. 1890. *The principles of psychology,* New York, Henry Holt.

JENKINS, W. O. & STANLEY, J. C. J. 1950. Partial reinforcement: a review and critique. *Psychological Bulletin,* 47**,** 193-234.

KAPUR, S., CRAIK, F. I. M., TULVING, E., WILSON, A. A., HOULE, S. & BROWN, G. M. 1994. Neuroanatomical correlates of encoding in episodic memory: levels of processing effect. *Proc Natl Acad Sci USA,* 91**,** 2008-2011.

KARTSOUNIS, L. D., RUDGE, P. & STEVENS, J. M. 1995. Bilateral lesions of CA1 and CA2 fields of the hippocampus are sufficient to cause severe amnesic syndrome in humans. *Journal of Neurology, Neurosurgery, and Psychiatry,* 59**,** 95-98.

KORSAKOFF, S. S. 1889. Etude médico-psychologique sur une forme des maladies de la mémoire. *Rev Philos,* 20**,** 501-530.

LEZAK, M. D., HOWIESON, D. & LORING, D. W. 2004. *Neuropsychological assessment (4th ed.),* Oxford, Oxford University Press.

LU, Z. L., WILLIAMSON, S. J. & KAUFMAN, L. 1992. Behavioral lifetime of human auditory sensory memory predicted by physiological measures. *Science,* 258**,** 1668-1670.

MAY, P. J. & TIITINEN, H. 2009. Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology,* (in press).

MILNER, B. 1962. Physiologie de l'hippocampe. *In:* PASSOUANT, P. (ed.). Paris: Centre National de la Recherche Scientifique.

NAATANEN, R., JACOBSEN, T. & WINKLER, I. 2005. Memory-based or afferent processes in mismatch negativity: a review of the evidence. *Psychophysiology,* 42**,** 25-32.

PAVLOV, I. P. 1927. *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex,* London, Oxford University Press.

PETERSON, L. R. & PETERSON, M. J. 1959. Short-term retention of individual verbal items. *Journal of Experimental Psychology,* 58**,** 193-198.

PEZDEK, K., WHETSTONE, T., KIRK, R., NUSHA, A. & THOMAS, D. 1989. Memory for real-world scenes: The role of consistency with schema expectation. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 15**,** 587-595.

PUTNAM, K. E. 1979. Hypnosis and distortions in eye witness memory. *International Journal of Clinical and Experimental Hypnosis,* 27**,** 437-448.

SAMS, M., HARI, R., RIF, J. & KNUUTILA, J. 1993. The human auditory sensory memory trace persists about 10 sec. *J Cog Neurosci,* 5**,** 363-370.

SCHNIDER, A. 2003. Spontaneous confabulation and the adaptation of thought to ongoing reality. *Nature Reviews Neuroscience,* 4**,** 662-671.

SCHNIDER, A., TREYER, V. & BUCK, A. 2000. Selection of currently relevant memories by the human posterior medial orbitofrontal cortex. *Journal of Neuroscience,* 20**,** 5880-5884.

SCOVILLE, W. B. & MILNER, B. 1957. Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry,* 20**,** 11-21.

SHIMAMURA, A. P., JANOWSKY, J. S. & SQUIRE, L. R. 1990. Memory for the temporal order of events in patients with frontal lobe lesions and amnesic patients. *Neuropsychologia,* 28**,** 803-813.

SIDEROFF, S., BUENO, O., HIRSCH, A., WEYAND, T. & MCGAUGH, J. 1974. Retrograde amnesia initiated by low-level stimulation of hippocampal cytoarchitectonic areas. *Exp Neurol,* 43**,** 285-297.

SQUIRE, L. R. 2009. The legacy of patient H.M. for neuroscience. *Neuron,* 61**,** 6-9.

STERNBERG, S. 1966. High-speed scanning in human memory. *Science,* 153**,** 652-654.

THOMPSON-SCHILL, S. L., JONIDES, J., MARSHUETZ, C., SMITH, E. E., D'ESPOSITO, M., KAN, I. P., KNIGHT, R. T. & SWICK, D. 2002. Effects of frontal lobe damage to interference effects in working memory. *Cognitive, Affective & Behavioral Neuroscience,* 2**,** 109-120.

ULANOVSKY, N., LAS, L., FARKAS, D. & NELKEN, I. 2004. Multiple time scales of adaptation in auditory cortex neurons. *J Neurosci,* 24**,** 10440-10453.

WICKENS, D. D., BORN, D. G. & ALLEN, C. K. 1963. Proactive inhibition and item similarity in short-term memory. *Journal of Verbal Learning and Verbal Behavior,* 2**,** 440-445.

8 Plasticity

The functional organization of the human brain—for instance that the temporal lobe processes auditory and occipital cortex visual information—is largely not hard-wired, unlike in some other species. Rather, the functional organization of the brain that we see in neuroimaging studies, or that which is reflected in the effects of focal brain damage, reflects the way that the brain has adapted to the environmental demands and tasks of the individual that for the most part stay fairly constant throughout the life. Indeed, it can be argued that functional organization of the brain is the end result of an adaptive process, reflecting all of the skills that one has learned to best cope with one's environment. Consequently, as long as the environment and the tasks and goals of the individual stay the same also the functional organization of the brain remains stable.

Upon changes in the environment, however, such as when moving to another country with entirely different culture and language, changes in the sensory inputs to the brain (that occurs when one of the senses is damaged such as blindness or deafness), or when the functional neuronal networks are themselves altered by for example stroke or some other type of brain damage, there is a sudden need for the brain to re-adapt. As the result of this, there are changes in the functional organization of the brain, with some of these changes occurring surprisingly quickly, in a matter of minutes, and some changes taking place over more extended periods of days, weeks, and months. In some cases these changes are extensive, and in some other cases they are rather subtle. These adaptive changes in functional brain organization are called plasticity. There are multiple forms of plasticity that occur in the brain that will be described in the following.

There are different forms of plasticity

Different forms of plasticity can be defined based on the reason why re-adaptation is required and by the developmental stage—whether it is a developing brain or an adult brain that is learning or re-learning. Indeed, the developing brain is far more plastic than the adult brain. Developmental plasticity refers to the rapid functional organization that occurs when the developing nervous system starts to adapt to it's environment, for instance, there are specific periods, called sensitive periods, during which acquiring certain skills such as language is easier than during other times.

Another class of plasticity studies have focused on characterizing robust reorganization in the functional architecture of the brain upon loosing sensory input due to, for instance, blindness, deafness, or loosing of a limb. This is also called activity-dependent plasticity, as the stimulus-related activity changes drastically. A third type of plasticity studies were introduced in Chapter 7 – perceptual or sensory learning, such as when learning to discriminate the phonemes of a foreign language, is made possible by very fine functional reorganization that takes place in the sensory cortices.

There is also a vast body of literature describing plasticity that follows damage to the neural networks themselves. Following brain damage, the challenge faced by the nervous system is to re-adapt with the remaining neural tissue to coping in the environment after previously learned skills and processing capacity are suddenly lost. In addition to these different types of long-lasting or even permanent plastic changes, rapid (or short-term) plasticity refers to functional re-organization that occurs very quickly, in a matter of seconds or minutes. As was already described in Chapter 6, short-term plasticity supports attentional filtering of task-relevant stimuli. What makes short-term plasticity especially interesting is that some of the rapid changes in neuronal tuning seem to linger, thus making perceptual/sensory learning possible. In the following, these different types of plasticity are described in more detail.

8.1 Developmental plasticity

The developing brain is highly plastic in humans. The downside of this is that the developing human is, compared to other species, practically helpless and without any skills required for survival for the first years of his/her life. Other species with multitude of innate abilities are able to function in their environment mere minutes or hours after birth. The advantage given by the high degree of brain plasticity is that humans are able to adapt to a far greater range of different environments than most of the other species. For instance, it is possible to learn any of the rather diverse languages and cultures of the world that one is born into, and learn diverse skills, from hunting to computer programming.

Sensitive/critical periods: windows of enhanced plasticity

There are specific sensitive/critical periods in development during which acquisition of certain skills is substantially easier than at other times. In the animal domain there are critical periods that are strictly timed. A good example of such is imprinting in birds; when a duckling is born, the first creature seen, even a human being, will be followed around like the mother duck. In humans, the prime example of a critical period is that of language learning; if a child is not exposed to language during the first few years of life, it will be very difficult (though not impossible (Krashen, 1973)) to acquire language later in life. The critical period for language learning was originally proposed by Eric Lenneberg. He hypothesized that lateralization of the brain (that at the time was thought to be completed after puberty) would be critical determinant for language learning ability (Lenneberg, 1967).

Subsequent studies on second language learning have challenged the critical period view. For instance, it has been shown that EEG responses during processing of an artificial language learned at adult age are similar to responses recorded to native language (Friederici et al., 2002) and robust second language learning has been observed at ages following brain lateralization (Snow and Hoefnagel-Höhle, 1978). Thus, rather than strict all-or-none critical periods, humans rather have sensitive periods during which acquisition of certain skills is significantly easier than earlier and later in life.

Besides language, other human cognitive functions also exhibit sensitive periods. As one example of this, development of social cognition (see Chapter 12) depends on exposure to social stimuli and social interactions during development. In a pioneering study autistic type of behavior was observed in monkeys that were reared in social isolation (Harlow et al., 1965). Autistic self-clutching behavior, seen in these monkeys when they are removed from social isolation, is shown in **Figure 8-1**. Monkeys isolated for 6 months showed significantly larger deficits in social behavior than those isolated for 3 months. In fact, there is accumulating evidence indicating that the reason why the human cortex (and especially the prefrontal cortex) has enlarged much more during phylogeny than in other species is due to the vast number and depth of social relationships that humans have to manage throughout the lifespan (Dunbar, 2009).

Figure 8-1. Shown is autistic self-clutching behavior exhibited by monkeys grown in social isolation when exposed to a social environment. Adapted from (Harlow et al., 1965).

In addition to higher cognitive functions, there are several other abilities/functions where there are sensitive periods early in development. From birth, the basic sensory processing abilities start to develop rapidly. While there are certain types of stimuli that infants prefer over others such as smiling faces (and even drawings resembling faces such as smileys) (Goren et al., 1975), the sensory statistics of the world that the developing individual acts in determines which sensory stimulus features will ultimately be represented in the brain.

Cats reared from birth to age of 10-12 weeks in such as fashion that they were never exposed to contrasts or lines of a certain orientation failed to develop visual cortex orientation columns that make it possible to detect those orientations (Hirsch and Spinelli, 1970), and to a large extent these effects lingered after prolonged exposure to normal visual stimulation (Spinelli et al., 1972). Similar effects were observed on the development of ocular dominance columns when monocular deprivation was used (Wiesel and Hubel, 1963). As will be described below, these experiments have provided clues on the neural mechanisms underlying plastic changes that occur during development.

Pruning of synaptic connections is important in development

In studies that have examined development of binocular rivalry, one of the eyes of the experimental animal was sutured shut for different periods during development. While no effects were observed if the blinding of the eye of a kitten took place prior to age of three weeks, or after the kitten was three months old, between the age of three weeks and three months the effect of depriving one eye of stimuli had a robust effect on the functional organization of the visual cortex. Specifically, the ocular dominance columns (see Chapter 5) failed to develop normally but rather the afferents from the non-deprived eye took over. This finding is explained by a fundamental phenomenon called pruning that helps shape the functional organization of the brain during development: the brain of children aged 1-2 years have significantly more synapses than the adult brain (Huttenlocher, 1979). Pruning of synapses that are not needed, given the environment of the individual, then contribute to learning of necessary skills as much as genesis of new synapses does—both effects serve the purpose of making the neural networks more organized and efficient in carrying out the information processing that is needed.

In a recent study, brain gray matter development, measured with MRI, was followed between ages of 4–21 years. It was observed that higher-order association cortices matured after lower-order sensory cortices and that phylogenetically older brain areas matured earlier than newer ones (Gogtay et al., 2004). These findings are illustrated in **Figure 8-2** below. What was specifically observed was loss of gray matter during development that could be attributed to a number of factors such as myelination, but also pruning of unnecessary synaptic connections with development. This cortical maturation pattern roughly matches the pattern of critical/sensitive periods for the development of perceptual and cognitive functions; sensory cortical processing that allows one to perceive sights and sounds develops first followed by higher-order cognitive functions.

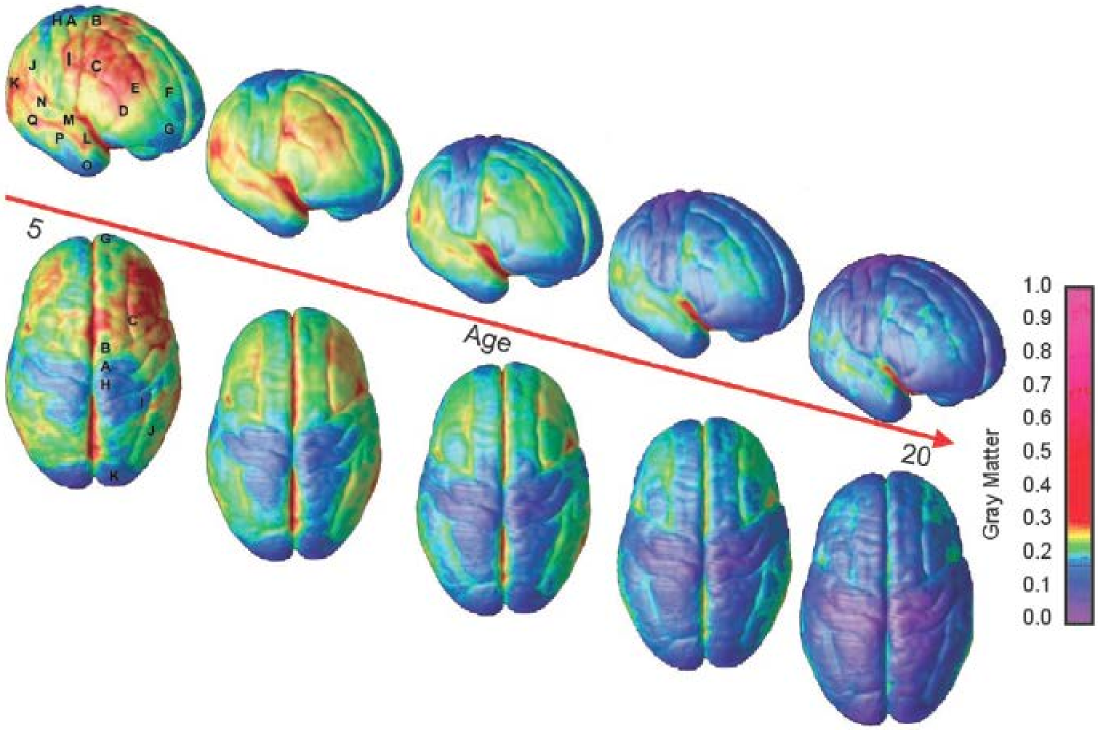


Figure 8-2: Shown is a sequence of gray matter loss during cortical development, measured with MRI, between ages 4–20 years. See text for details. Adapted from (Gogtay et al., 2004).

Neurons that fire together wire together

Importantly, the firing patterns of neurons are what determine which of the synapses are discarded and which are strengthened. This principle has been often referred to as "neurons that fire together, wire together", as temporally correlated firing caused by, for instance, simultaneous inputs that arrive from the same eye, strengthens synapses and those synapses that do not receive temporally correlated firing die out (see **Figure 8-2**). This principle is also called Hebbian learning, after Donald Hebb who first proposed the mechanism in the late 1940s "*When an axon of cell A is near enough to excite cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased*" (Hebb, 1949).

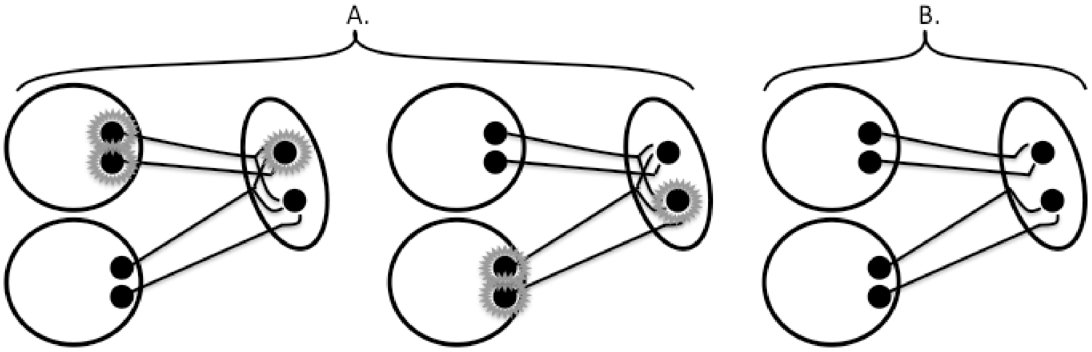


Figure 8-3. Schematic illustration of the principle "Neurons that fire together, wire together". A. Neurons of a given eye fire in temporal synchrony onto neurons in lateral geniculate nucleus of the thalamus, whereas neurons from different eyes do not. Initially, in the developing nervous system, there are also synaptic connections from the other eye to the lateral geniculate nucleus neurons. B. As a result of synchronous firing, the connections from one of the eyes have survived and become stronger, whereas those connections that failed to exhibit temporally synchronous firing have been pruned. This is a general developmental principle as an infant has more synaptic connections than an adult and in the course of development of perceptual and cognitive skills the connections that are not used are removed. Pruning is believed to be an even more important mechanism than activity-dependent genesis of new synaptic connections.

Synchronous inputs strengthen synapses *via* NMDA-gated Ca^{2+} channels

The strengthening of synaptic connections during temporally congruent firing is made possible by N-methyl-D-aspartate (NMDA) receptor gated calcium ion channels. When the degree of presynaptic stimulation is low (*i.e.*, when there are few simultaneous inputs that summate) the channels are blocked by magnesium ions and inflow of calcium is low. But when the membrane potential depolarizes beyond a certain threshold due to multiple simultaneous excitatory post-synaptic potentials (caused by action potentials of pre-synaptic neurons) the magnesium ions are displaced from the NMDA-gated channels and there is sudden flow of calcium into the intracellular space (Nowak et al., 1984). This depolarizes the neurons, and results in strengthening of the synaptic connections (Markman et al., 1997). These mechanisms, also known as long-term potentiation, were first observed in the hippocampus, and have been associated with formation of new memories, as was briefly described in the preceding chapter of this book.

The developing brain recovers from injury better than the adult brain

The high degree of plasticity in the developing brain is also reflected in how the brain recovers from injuries. Recovery from injuries is both faster and more complete in children than in adults. A striking example of this is offered by structural neuroimaging findings in persons who are suffering from hydrocephalus. Hydrocephalus is a condition where intracranial pressure is abnormally high and cerebrospinal fluid accumulates and gradually displaces brain matter. As a result of this, ventricles containing cerebrospinal fluid are enlarged and in some cases there is very little cortex left at adult age. Despite this, however, these persons can have even university degrees and live normal social lives (Lorber, 1981, Lewin, 1980). An example of this type of person's brain is shown in **Figure 8-4**.

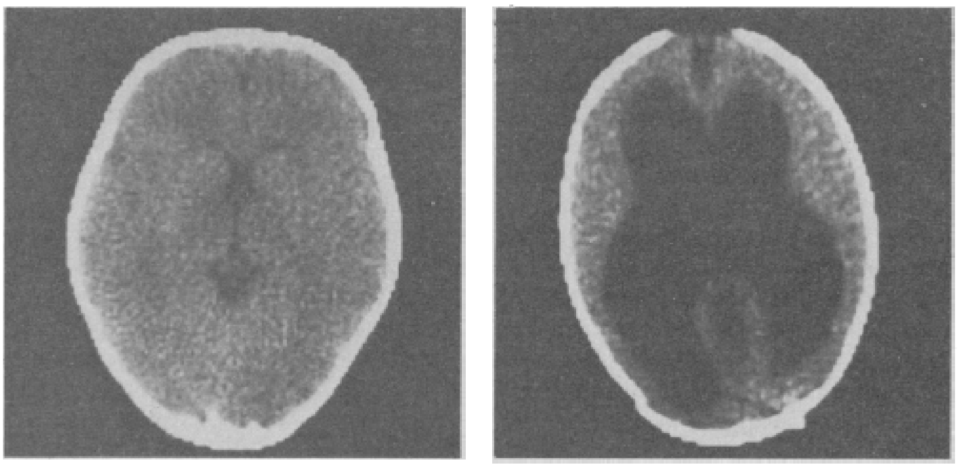


Figure 8-4. Axial scans of normal (left) and hydrocephalic (right) brain, showing how little there is cortex left as the result of abnormally high cerebrospinal fluid pressure enlarging the ventricles over the course of years. Yet such patients, when the condition develops gradually from childhood, can exhibit above normal intelligence and be fully functional in other respects too, demonstrating the extent that there is plasticity in the developing brain. Adapted from (Lewin, 1980).

In these cases, it is assumed that the slow development of the condition throughout childhood and adolescence allows the brain to adjust to the changes, and thus attain relatively normal perceptual and cognitive skills despite the gross anatomical changes in brain structure. The same is true for other types of brain injuries that occur during childhood; brain damage of an extent that would leave an adult crippled can exhibit little effect in later life when suffered in childhood. One possible explanation for this higher extent of plasticity is that there are more synapses at the young age that can be pruned.

8.2 Plasticity due to loss of sensory input

Reorganization of brain functions following alteration in sensory input or loss of an entire sense provides another demonstration of brain plasticity. As was described in Chapter 5, a surprisingly large part of the cortex is devoted to processing of sensory inputs, with superior temporal lobe areas processing auditory, occipital areas visual, and anterior parietal cortical areas somatosensory information. It has therefore been a puzzling question what happens to sensory cortex when it no longer receives sensory input due to blindness or deafness, and whether there are differences between early and late age of onset of the sensory loss. The question that early neuroimaging studies set forth to answer was whether the sensory-deprived cortex deteriorates or whether it begins to process other inputs.

Sensory cortex processes inputs of preserved modalities in blind and deaf

When the visual cortex of blind persons has been investigated with modern functional imaging methods, it has become obvious that the visual cortex has not deteriorated due to lack of visual input (Phelps et al., 1981), as one might predict based on the "neurons that fire together wire together" principle. In contrast, the metabolic activity exhibited by the visual cortex of the blind equaled that of normal-sighted persons at rest (Phelps et al., 1981) and, in subsequent studies, it was observed that the visual cortex in blind is processing auditory stimuli (Kujala et al., 1995), which is illustrated in **Figure 8-5**.

This observation provides a potential explanation why certain perceptual abilities in blind are enhanced such as ability to localize sounds in space based on hearing alone, and/or enhanced somatosensory discrimination such that is needed, for instance, when reading Braille or feeling textures or shapes of objects. Supporting this view, activation of the visual cortex in blind subjects was observed during Braille reading, but not during somatosensory stimulation that did not require fine discrimination ability (Sadato et al., 1996).

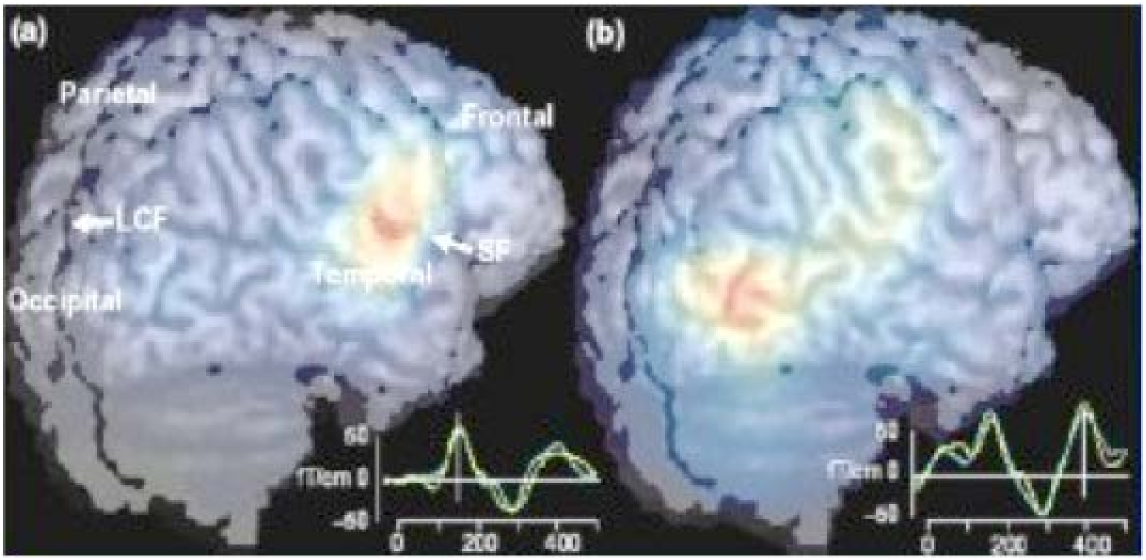


Figure 8-5: Magnetoencephalography shows processing of auditory stimuli at a relatively early latency of ~200 ms in the visual cortex of the blind perons. a) Healthy volunteer exhibits focal responses in the auditory cortex (the inferior frontal responses are possibly crosstalk from the nearby auditory areas), whereas b) robust occipital cortical responses are evident in the blind person. Adapted from (Kujala et al., 1995).

While the observations of functional re-organization were early on limited to congenitally blind, it was subsequently observed that similar, although less robust, functional reorganization takes place also in persons who become blind later in life, in adolescence and even adulthood, well beyond the period of robust plasticity of infancy/childhood (Kujala et al., 1997). Furthermore, in addition of observations of visual cortex processing auditory and somatosensory stimuli in the blind, analogous functional reorganization has been observed in deaf persons. In the deaf, it is the auditory cortex that has been observed to process inputs from other sense modalities (Finney et al., 2001).

Cross-modal connections in sensory cortical areas

Cross-modal connectivity is one potential explanation for the cross-modal plasticity observed in blind and deaf persons. It is known based on both anatomical and functional imaging studies that the sensory cortices are not as unisensory as once believed. There are direct anatomical connections even between primary auditory and visual cortices (Falchier et al., 2002). Visual stimuli activate the auditory cortex, and auditory stimuli elicit responses in the visual cortex, at very early latencies of a few tens of milliseconds (Raij et al., 2010). Thus, when a sense is lost together with its inputs to a given cortical area, there are connections from other senses and brain areas that continue to stream in to the area. Synchronous inputs then begin to shape synaptic connectivity, resulting in emergence of functionality such as enhanced somatosensory discrimination ability that supports Braille reading.

Another quite striking finding of cortical plasticity due to change in sensory stimulation comes from studies where the auditory inputs to the auditory cortex were first cut off and then replaced with visual inputs in developing ferrets (Sur et al., 1988). This re-direction of visual input to auditory cortex resulted in emergence of neuronal receptive fields in the auditory cortex that resembled visual simple and complex cells (see Chapter 5 for a description of simple and complex cells). The animals were also able to navigate in their environment based on visual stimuli, suggesting that despite subtle differences in microanatomy between different cortical areas, the input statistics is what shapes the functional organization of the cortex, rather than there being some innate properties that determine functionality of each cortical area.

Plasticity can take place relatively quickly

While major reorganization probably takes place during weeks, months, or even years, some of the plasticity changes take place surprisingly quickly. In animal studies, it has been observed that somatosensory cortical areas innervating a given digit (in case of monkey studies, or a whisker in case of rats) are taken over by inputs from adjacent digits/whiskers (Merzenich et al., 1984). When local anesthesia has been used to deplete inputs from a given digit to the somatosensory cortex, signs of reversible functional reorganization have been documented to occur immediately from the onset of local anesthesia (Calford and Tweedale, 1991).

These findings have been attributed mainly to unmasking type of phenomena (*i.e.*, release from inhibition exhibited by incoming somatosensory inputs). However, even though the inputs take over very quickly after a change in sensory input, in order for the reorganized representations to become functionally meaningful/useful, such as enhanced spatial hearing in blind persons, extensive training (*i.e.*, exposure and functioning in the everyday environment with the altered sensory input) is required over weeks, months, or even years. This probably involves both genesis and strengthening of existing synapses and discarding of unused ones that together shapes the networks of neurons to support novel functionality.

8.3 Smaller scale plasticity: sensory/perceptual learning

In addition to large-scale functional reorganization that takes place when a sense is lost, there are more subtle changes in functional organization of sensory cortices when one learns to discriminate stimuli from one another. Consider, for instance, the case of learning a foreign language. There are specific phonetic sounds in foreign languages that are not contained in one's mother tongue (for example "r" and "l" of English are not discriminated by Japanese speakers). Then, in order to fully master a given foreign language, one needs to learn to discriminate the previously unknown phonemes.

Another example where one needs to learn to discriminate between specific sounds is dissociating specific instrument timbres from each other, for instance, those of violin *vs*. viola. Modern neuroimaging methods enable study of plasticity underlying this type of perceptual learning. For instance, auditory cortex responses specific to native language phonemes have been documented (Naatanen et al., 1997), and recording such responses prior to and in the course of language learning allows one to track plastic changes in the brain associated with perceptual learning.

Findings from both human and animal studies suggest that sensory learning is made possible by changes in receptive fields of auditory cortical neurons (for recent reviews, see (Jaaskelainen et al., 2007, Ahissar et al., 2009)). During sensory discrimination learning, receptive fields of a portion of neurons are retuned to encompass the critical features of the to-be-discriminated stimulus. In rats, task-specific auditory cortex tonotopic map changes were observed when rats learned to discriminate target sounds of specific frequency and corresponding changes in auditory cortex ampliotopic maps were observed when the animals learned to discriminate sounds of specific intensity (Polley et al., 2006).

Human neuroimaging studies have shown that musical training enlarges auditory cortex representations of sound features relevant for music perception such as the timbre of the instrument that the musician plays (Pantev et al., 2003). It has been further suggested that selective attention to the to-be-learned stimulus features is essential for sensory learning in adults (Ahissar et al., 2009), but not in infancy/childhood where the functional organization appears to take place due to stimulus statistics. Acetylcholinergic input to cortex from the nucleus basalis of Meynert is what appears to make it possible for longer lasting, or even permanent, receptive field changes to take place (Froemke et al., 2007).

8.4 Plasticity associated with recovery from brain damage

Brain damage poses a great challenge for brain plasticity. Following brain damage, the damaged functional systems need to be reorganized in order to recover lost abilities and/or develop compensatory/alternative abilities. This challenge varies with the type of brain damage. The quicker the brain damage occurs the larger the deficits that even minor lesions can cause. Naturally, the site of the lesion dictates in which ways and how severely impaired the afflicted person will become. For instance, damage to brainstem that supports vital functions can be life-threatening even if the lesion is relatively small. On the other hand, damage to frontal lobes can even alter one's personality.

Slowly developing brain damage

The hydrocephalus, which was described above, and brain tumors are good examples of brain damage that develop at a slow pace. The damage caused by tumors depends on the type of tumor. Benign tumors cause displacement of brain tissue by occupying increasingly large portion of the intracranial space. In case of malign tumors, the cancerous tissue infiltrates into the brain tissue. This makes it more difficult to surgically remove malign tumors. The relatively slow pace at which brain tissue is damaged by tumors often results, much like in the case of hydrocephalus described above, in relatively minor cognitive and perceptual deficits early on.

In fact, it is often the case that brain tumors are not detected until they are already quite large at which stage the increased intracranial pressure causes persistent headache and nausea. Then, removal of the brain tumor is bound to also remove larger portion of healthy brain tissue. This causes sudden changes to the existing functional organization of the brain and an additional challenge to brain plasticity. Naturally, development of more advanced and novel cancer treatment methods such as cancer vaccines, medications that prevent angiogenesis that is vital for cancerous tissue, and increased accuracy in surgical methods (that are exemplified by development of highly accurate stereotactic radiotherapy methods such as the Gamma Knife) are likely to diminish this problem in the future.

Other cases of slowly occurring brain damage are various types of dementia. Some dementias are manifested with relatively mild cognitive deficits even though there might already be extensive damage to the underlying neural systems. As an example of this, clinical symptoms of Parkinson's disease do not take place until the vast majority (70-95%) of dopaminergic neurons that ascend from substantia nigra to striatum have died. The most common form of dementia is the Alzheimer's disease that is manifested initially by memory problems (that the patients seem unaware of) followed by quick deterioration of cognitive functions. Another quite common cause of dementia is vascular dementia, for instance, in multi-infarct dementia there are, due to high blood pressure, multiple small infarcts that progressively lead to deterioration of cognitive functions.

Sudden-onset brain damage

Sudden-onset brain damage can occur for a variety of reasons. A sudden impact when hitting the head in, for instance, a car accident can cause sudden brain damage. The brain tissue under the part of the skull that takes the hit typically shows focal damage, however, even larger damage can often be seen in the opposite side of the brain. This is due to the impact sending the brain in motion and pressing the brain against the skull on the opposite side, a phenomenon that has been termed coup counter. Furthermore, the sudden twisting of the brain often cuts axonal connections that can result in severe cognitive deficits without visible damage in conventional clinical MRI.

Another form of sudden brain damage results from lack of oxygen supply, anoxia, which occurs, for instance, when drowning or during carbon monoxide poisoning (carbon monoxide displaces oxygen in the red blood cells). Focal anoxia, on the other hand, results from failures in blood supply (collectively called stroke) such as when a clot stuck in one of the arteries blocks circulation to specific brain areas, or when one of the arteries is ruptured (a condition called hemorrage). In the latter case, accumulation of blood in the intracranial space further results in increased intracranial pressure and surgical operation is needed to alleviate the pressure as well as to repair the rupture.

In each of these cases of sudden onset brain damage, the cognitive and perceptual deficits are initially quite severe and they are further enhanced by brain edema that subsides in the course of the following days and weeks. It is quite typical that there is very rapid “spontaneous recovery” of cognitive functions during the first three months after the damage has occurred. It was not until quite recently that clinicians thought that there is little possibility for further recovery following the period of spontaneous recovery, however, increased recognition of the extent of plasticity that there is even the adult brain has led to increased optimism towards rehabilitation of cognitive functions. Perhaps the most important principle in treatment of brain damage is, however, trying to avoid re-occurrence; for instance, in case of stroke, one should try to avoid/remove any predisposing factors such as high blood pressure. Reoccurrence of the stroke typically results in enhanced severity of cognitive deficits.

Rehabilitation of cognitive functions following brain damage

There are basically two approaches to rehabilitation. The first one is learning of compensatory skills and alternative actions to the damaged one, for instance, use of external memorizing aids when suffering from memory deficits. The second type of rehabilitation is targeted training of the damaged function. One example of this latter type is rehabilitation of a scotoma, loss of vision in a portion of ones field of view. Showing a light at the edge of a scotoma in an active detection task in patients suffering from geniculostriate damage resulted in gradual shrinking of the extent of scotoma (Zihl and von Cramon, 1982).

This is similar to the sensory/perceptual learning tasks described above that have been associated with alterations in the receptive fields of sensory cortical neurons. As such it is a positive example of the possibilities offered by rehabilitation even in the adult brain. From the perspective of rehabilitation the most challenging cases of brain damage are patients who lack insight into their cognitive and other deficits. This is often the case with prefrontal cortical damage. In these patients, it is important to give timely feedback when they behave in an impulsive and inappropriate manner, to gradually increase their insight.

8.5 Neurogenesis, stem cell treatments and plasticity

For a long time it was a widely held belief that new neurons are not generated after birth, even though there was experimental evidence tentatively suggesting that neurogenesis does take place in the brain published already in the early 20th Century (Allen, 1912) and more substantive evidence was obtained in studies conducted in the 1960s (Altman, 1962). Subsequently, it has become obvious that neurogenesis does happen. The extent/rate of neurogenesis depends on the brain area (for a recent review, see (Migaud et al., 2010)). In hippocampus it is very frequent, whereas in prefrontal cortical areas there seems to be little neurogenesis. Neurogenesis does appear to be relevant for plasticity, learning and memory, as elevated levels of stress (due to elevated levels of stress hormones such as cortisol) lead to significant reduction in the rate of neurogenesis, and it is well known that memory deficits take place under conditions of elevated stress (for a recent review on the effects of stress on hippocampal neurogenesis, see (Dranovsky and Hen, 2006)).

The arrest of neurogenesis might also be associated with gray matter volume reduction in hippocampus that has been observed in association with prolonged elevated stress such that occurs in posttraumatic stress disorder (Kitayama et al., 2005). However, given that the presence of neurogenesis in the adult brain was acknowledged relatively recently, the precise roles played by this phenomenon in cognitive functions, including plasticity, are not well known and are currently a topic of intense research. Neurogenesis cannot, however, replace dead neurons in extensive brain areas such as those destroyed by a stroke, since the glia cells quickly remove the dead neural tissue and grow in place to replace it. However, isolation of adult stem cells from tissue samples of the patient him/herself and injection of them to replace lost dopaminerigic neurons has been reported to yield very promising results in treatment of idiopathic Parkinson's disease (Lévesque et al., 2009). This has the advantage that, unlike when embryonic stem cells are utilized, there is no immune reaction against the adult stem cells.

8.6 Concluding remarks

Over the past few decades, impressive demonstrations of the brain's capacity for functional re-organization have followed one after another. It is now known that the sensory and higher-order cognitive systems of the brain exhibit across-time stable functional anatomy more due to the environment, stimuli, tasks, and goals of the individual remaining fairly stable, rather than due to fixed nature of the central nervous system, as was once thought. While the developing brain is by far more plastic than the adult brain, there is still substantial capacity for plasticity in the latter.

Learning happens both as a result of pruning of unnecessary synapses as well as generation of new ones and strengthening of the central ones that consistently receive simultaneous inputs from other neurons. It is also now known that there is extensive neurogenesis in the brain, especially in hippocampus that probably relates to the role of this structure in memorizing and learning that was described in Chapter 7. Overall, findings on brain plasticity are highly important in that they make it possible to develop various rehabilitation methods for those suffering from various types of brain damage.

8.7 References

AHISSAR, M., NAHUM, M., NELKEN, I. & HOCHSTEIN, S. 2009. Reverse hierarchies and sensory learning. *Philos Trans R Soc Lond B Biol Sci,* 364, 285-299.

ALLEN, E. 1912. The cessation of mitosis in the central nervous system of the albino rat. *Journal of Comparative Neurology,* 22, 547-568.

ALTMAN, J. 1962. Are new neurons formed in the brains of adult mammals? *Science,* 135, 1127-1128.

CALFORD, M. B. & TWEEDALE, R. 1991. Immediate expansion of receptive fields of neurons in area 3b of macaque monkeys after digit denervation. *Somatosensory & Motor Research,* 8, 249-260.

DRANOVSKY, A. & HEN, R. 2006. Hippocampal neurogenesis: regulation by stress and antidepressants. *Biol Psychiatry,* 59, 1136-1143.

DUNBAR, R. I. 2009. Darwin and the ghost of Phineas Cage: neuro-evolution and the social brain. *Cortex,* 45, 1119-1125.

FALCHIER, A., CLAVAGNIER, S., BARONE, P. & KENNEDY, H. 2002. Anatomical evidence of multimodal integration in primate striatal cortex. *J Neurosci,* 22, 5749-5759.

FINNEY, E. M., FINE, I. & DOBKINS, K. R. 2001. Visual stimuli activate auditory cortex in the deaf. *Nature Neuroscience,* 4, 1171-1173.

FRIEDERICI, A. D., STEINHAUER, K. & PFEIFER, E. 2002. Brain signatures of artificial language processing: evidence challenging the critical period hypothesis. *Proc Natl Acad Sci USA,* 99, 529-534.

FROEMKE, R. C., MERZENICH, M. M. & SCHREINER, C. E. 2007. A synaptic memory trace for cortical receptive field plasticity. *Nature,* 450, 425-429.

GOGTAY, N., GIEDD, J. N., LUSK, L., HAYASHI, K. M., GREENSTEIN, D., VAITUZIS, A. C., NUGENT, T. F., HERMAN, D. H., CLASEN, L. S., TOGA, A. W., RAPOPORT, J. L. & THOMPSON, P. M. 2004. Dynamic mapping of human cortical development during childhood through early adulthood. *Proc Natl Acad Sci USA,* 101**,** 8174-8179.

GOREN, C. C., SARTY, M. & WU, P. Y. 1975. Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics,* 56**,** 544-549.

HARLOW, H. F., DODSWORTH, R. O. & HARLOW, M. K. 1965. Total social isolation in monkeys. *Proc Natl Acad Sci USA,* 54**,** 90-97.

HEBB, D. O. 1949. *The organization of behavior,* New York, Wiley.

HIRSCH, H. V. B. & SPINELLI, D. N. 1970. Visual experience modifies distribution of horizontally and vertically oriented receptive fields in cats. *Science,* 168**,** 869-871.

HUTTENLOCHER, P. R. 1979. Synaptic density in human frontal cortex - developmental changes and effects of aging. *Brain Research,* 163**,** 195-205.

JAASKELAINEN, I. P., AHVENINEN, J., BELLIVEAU, J. W., RAIJ, T. & SAMS, M. 2007. Short-term plasticity in auditory cognition. *Trends in Neurosciences,* 30**,** 653-661.

KITAYAMA, N., VACCARINO, V., KUTNER, M., WEISS, P. & BREMMER, J. D. 2005. Magnetic resonance imaging (MRI) measurement of hippocampal volume in posttraumatic stress disorder: a meta-analysis. *Journal of Affective Disorders,* 88**,** 79-86.

KRASHEN, S. D. 1973. Lateralization, language learning, and the critical period: some new evidence. *Language learning,* 23**,** 63-74.

KUJALA, T., ALHO, K., HUOTILAINEN, M., ILMONIEMI, R. J., LEHTOKOSKI, A., LEINONEN, A., RINNE, T., SALONEN, O., SINKKONEN, J., STANDERTSKJÖLD-NORDENSTAM, C.-G. & NÄÄTÄNEN, R. 1997. Electrophysiological evidence for cross-modal plasticity in humans with early- and late-onset blindness. *Psychophysiology,* 34**,** 213-216.

KUJALA, T., HUOTILAINEN, M., SINKKONEN, J., AHONEN, A. I., ALHO, K., HÄMÄLÄINEN, M. S., ILMONIEMI, R. J., KAJOLA, M., KNUUTILA, J. E. T., LAVIKAINEN, J., SALONEN, O., SIMOLA, J., STANDERTSKJÖLD-NORDENSTAM, C.-G., TIITINEN, H., TISSARI, S. & NÄÄTÄNEN, R. 1995. Visual cortex activation in blind humans during sound discrimination. *Neuroscience Letters,* 183**,** 143-146.

LENNEBERG, E. H. 1967. *Biological Foundations of Language,* New York, John Wiley & Sons.

LÉVESQUE, M. F., NEUMAN, T. & REZAK, M. 2009. Therapeutic microinjection of autologous adult human neural stem cells and differentiated neurons for Parkinson's disease: five-year post-operative outcome. *Open Stem Cell Journal,* 1**,** 20-29.

LEWIN, R. 1980. Is your brain really necessary. *Science,* 210, 1232-1234.

LORBER, J. 1981. Is your brain really necessary? *Nurs Mirror,* 152, 29-30.

MARKMAN, H., LÜBKE, J., FROTSCHER, M. & SAKMANN, B. 1997. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science,* 275, 213-215.

MERZENICH, M. M., NELSON, R. J., STRYKER, M. P., CYNADER, M. S., SCHOPPMANN, A. & ZOOK, J. M. 1984. Somatosensory cortical map changes following digit amputation in adult monkeys. *Journal of Comparative Neurology,* 224, 591-605.

MIGAUD, M., BATAILLER, M., SEGURA, S., DUITTOZ, A., FRANCESCHINI, I. & PILLON, D. 2010. Emerging new sites for adult neurogenesis in the mammalian brain: a comparative study between the hypothalamus and the classical neurogenic zones. *European Journal of Neuroscience,* 32, 2042-2052.

NAATANEN, R., LEHTOKOSKI, A., LENNES, M., CHEOUR, M., HUOTILAINEN, M., IIVONEN, A., VAINIO, M., ALKU, P., ILMONIEMI, R. J., LUUK, A., ALLIK, J., SINKKONEN, J. & ALHO, K. 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature,* 385, 432-434.

NOWAK, L., BREGESTOWSKI, P., ASCHER, P., HERBERT, A. & PROCHIANTZ, A. 1984. Magnesium gates glutamate-activated channels in mouse central neurones. *Nature,* 307, 462-465.

PANTEV, C., ROSS, B., FUJIOKA, T., TRAINOR, L. J., SCHULTE, M. & SCHULZ, M. 2003. Music and learning-induced cortical plasticity. *Ann N Y Acad Sci,* 999**,** 438-450.

PHELPS, M. E., MAZZIOTTA, J. C., KUHL, D. E., NUWER, M., PACKWOOD, J., METTER, J. & ENGEL, J. J. 1981. Tomographic mapping of human cerebral metabolism visual stimulation and deprivation. *Neurology,* 31**,** 517-529.

POLLEY, D. B., STEINBERG, E. E. & MERZENICH, M. M. 2006. Perceptual learning directs auditory cortical map reorganization through top-down influences. *Journal of Neuroscience,* 26**,** 4970-4982.

RAIJ, T., AHVENINEN, J., LIN, F.-H., WITZEL, T., JAASKELAINEN, I. P., LETHAM, B., ISRAELI, E., SAHYOUN, C., VASIOS, C., STUFFLEBEAM, S., HAMALAINEN, M. & BELLIVEAU, J. W. 2010. Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices. *Eur J Neurosci,* 31**,** 1772-1782.

SADATO, N., PASCUAL-LEONE, A., GRAFMAN, J., IBANEZ, V., DEIBER, M.-P., DOLD, G. & HALLETT, M. 1996. Activation of the primary visual cortex by Braille reading in blind subjects. *Nature,* 380**,** 526-528.

SNOW, C. E. & HOEFNAGEL-HÖHLE, M. 1978. The critical period for language acquisition: evidence from second language learning. *Child Development,* 49**,** 1114-1128.

SPINELLI, D. N., HIRSCH, H. V. B., PHELPS, R. W. & METZLER, J. 1972. Visual experience as a determinant of the response characteristics of cortical receptive fields in cats. *Experimental Brain Research,* 15**,** 289-304.

SUR, M., GARRAGHTY, P. E. & ROE, A. W. 1988. Experimentally induced visual projections into auditory thalamus and cortex. *Science,* 242**,** 1437-1441.

WIESEL, T. N. & HUBEL, D. H. 1963. Single-cell responses in striate cortex of kittens deprived of vision in one eye. *Journal of Neurophysiology,* 26**,** 1003-1017.

ZIHL, J. & VON CRAMON, D. 1982. Restitution of visual field in patients with damage to the geniculostriate visual pathway. *Hum Neurobiol,* 1**,** 5-8.

9 Speech and Language

9.1 Language is a multifaceted and complex cognitive function

Imagine a situation where one is in a foreign country and cannot utter a word in the local tongue. Further, imagine that gesturing in that country would follow a peculiar logic of its own that one could not comprehend. Under such circumstances one quickly comes to realize the importance of being able to communicate with others, through speaking, gesturing, reading, and/or writing. Indeed, the seeming ease at which one normally commands one's mother tongue is highly misleading. In reality, language is one of the most complex and amazing of human cognitive functions.

Difficulties in developing computer programs that would be able to command language well enough to carry a conversation with a human is one example of the complexity of the task that the brain faces when listening and speaking. Further, out of all species, mankind is the only one capable of the kind of highly sophisticated and abstract form of communication that language represents. To some extent it has been debated whether language is unique to humans (*i.e.*, whether ”speech is special”) and there are language-like communicative calls and gesturing in other animals. The most striking manifestation of language ability in other species have been obtained in studies where sign language was taught to non-human primates (Gardner and Gardner, 1969), however, even the best of these examples fail to compare with the richness of human language.

Human language is a multifaceted skill. In order to communicate effectively, one has to know the meanings of different words (semantic knowledge) and have sufficient command of grammar to organize the words in an appropriate manner to be understood. Further, in order to read and write, one also has to have command of the orthographic (how words are written) and phonological (how words are pronounced) knowledge. With respect to this, there is variability between different languages, with prominent ambiguity in English (as there is no one-to-one correspondence between the written form and pronunciation) and non-ambiguous pronunciation in some other languages such as Italian and Finnish, where there is one-to-one correspondence between the written form and pronunciation.

9.2 Behavioral studies of human language

Today, there is a wealth of information about human language and the underlying neural mechanisms. A lot of that information has been accumulated in studies on language deficits exhibited by neurological patients with focal brain lesions and, more recently, in non-invasive neuroimaging studies with healthy volunteers. Prior to development of modern non-invasive neuroimaging methods, however, behavioral experiments in healthy volunteers have been carried out in a rather ingenious manner to reveal important aspects of human language functions. Examples of how behavioral methods have been utilized in human language studies are described in the following.

Mental lexicon

The vocabulary of the human language (also called “mental lexicon”) is highly excessive with, depending on the language, the amount of words ranging from tens of thousands to hundreds of thousands. The relative size of the vocabulary depends on how one counts and the nature of the language in question, for instance, whether only root words are counted (in which case so-called isolating languages appear to be larger than other languages) or whether one counts only words that are currently used. Nonetheless, vocabularies of each of the human languages are extremely rich, yet one is able to access each word relatively effortlessly, given that the words are produced and comprehended with an approximate rate of three per second (there is, of course, variability between languages in the average length of a single word).

The organization of the mental lexicon has been a topic of numerous behavioral studies (as well as that of neuroimaging studies, as will be described later in this chapter). As a simple example of how recording of reaction times could be used to study mental lexicon is measurement of reaction times to words that begin with different letters of the alphabet to answer the question of whether the mental lexicon is organized in an alphabetical order like any ordinary dictionary. If this were true, one would expect longer reaction times to words as a function of serial position within the alphabetically organized dictionary (*i.e.*, longer reaction times for words that begin with letter "w" than for words beginning with letter "a"). In reality, of course, the mental lexicon is not organized in an alphabetical order.

Rather than alphabetic organization, it has been observed that words which are more frequently used in daily life are accessed faster than less frequently used words (Forster and Chambers, 1973). Incidentally, it also seems that the more frequently used words are the ones that evolve slowly. For instance, across European languages the words for "two" are rather similar. In contrast, the words for "tail" are rather different (Pagel et al., 2007). In addition to frequent words, words that are more unique in how they sound are accessed faster than words that are acoustically less unique (*e.g.*, words "eight", "late", and "rate" sound alike; homophones such as "too" and "two" are the prime example of this).

Semantic networks of words

Behavioral studies have suggested that words which are semantically close to one another also seem to be close to one another at the level of brain representations. These so-called priming studies are based on a relatively simple experimental design where two words are presented in close succession. The first word is a prime and the second word is a test word. The test word is either a real word or a pseudoword (*i.e.*, a string of letters that resembles a real word but does not mean anything such as "tacke"). The task of subjects is then to press, as quickly as possible, a reaction time button if they judge the test word to be a real word. If the first word is semantically close to the test word, the reaction times are significantly faster than if the two are semantically unrelated (Neely, 1977). For instance, truck-car pairing results in faster responses than rose-car pairing.

Based on results from these studies, it has been possible to construct semantic networks of words (see **Figure 9-1**), where semantic concepts are nodes of the network and there are links between the nodes representing semantic associations between the nodes. The distance in the network (*i.e.*, number of links between two concepts) reflects distance in the semantic space. It is assumed that presentation of a word or a concept results in spreading of activation across the semantic network, thus explaining the priming effects. For an early review on the behavioral studies on semantic language networks, see (Collins and Loftus, 1975).

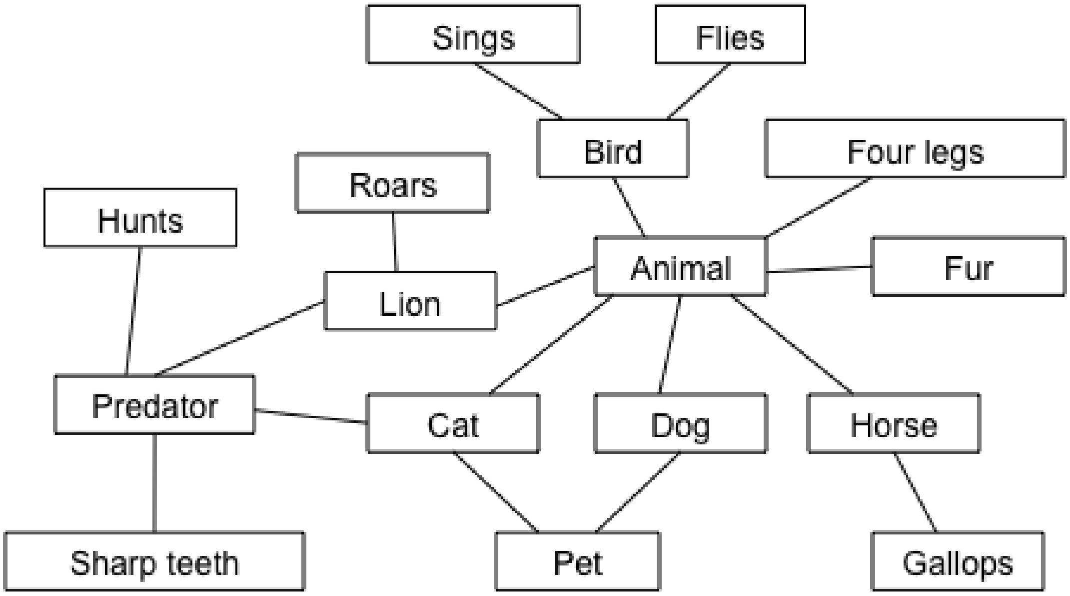


Figure 9-1: An example of a semantic network, based on reaction time studies, where semantically related prime word speeds up reaction time to a subsequent word that is semantically related.

9.3 Neural basis of speech and language

Elucidating the neural basis of speech and language has been one of the greatest challenges in cognitive neuroscience. Much of what we know today about the neural basis of language is based on careful studies of language disturbances with neuropsychological test batteries in patients who have suffered circumscribed brain lesions due to neurological conditions. For instance, it has been shown in patient studies that the various language skills, speaking, comprehending, reading, and writing are governed by partly distinct underlying neural mechanisms; each of these abilities can be separately severed following brain damage. For instance, there are brain-damaged patients who can write their thoughts on a piece of paper but are not able to read what they themselves just wrote. During the last few decades it has been possible to investigate the underlying neural mechanisms in healthy subjects using non-invasive brain imaging. In the following, key findings from patient and neuroimaging studies will be described.

Early studies in neurological patients: Broca and Wernicke

As briefly described in Chapter 2, Paul Broca and Carl Wernicke were the first to document brain-language interrelationships in patients who had language deficits following brain damage. The French physician Broca was the first one; he described a patient who could only utter a single word "tan" following brain damage (Broca, 1861). Curiously, his patient was still able to comprehend speech even though he had lost his ability to speak. After the patient later deceased, Broca conducted an autopsy to reveal which parts of the brain had been damaged. The area implicated in this early study, roughly encompassing the pars triangularis and pars opecularis of the inferior frontal gyrus on the left side, has widely become known as the Broca's area.

Following this early observation, the type of aphasia where speech production is severed is called Broca's aphasia (aphasia meaning "no speech", and dysphasia referring to a milder condition and/or one that concerns the developing brain). In subsequent patient studies, damage to Broca's area has also been associated with disturbances in the ability to produce grammatically correct sentences. These patients might, for instance, omit the "It is" from the sentence "It is dark outside", uttering only "dark outside". Furthermore, even when they are to repeat the sentence when it is presented to them, these patients might fail to repeat the sentence correctly, uttering only "dark outside".

Shortly after the observations of Broca, Austrian physician Carl Wernicke reported a brain lesion in the left temporal lobe in his patient that was associated with inability to comprehend speech (Wernicke, 1874). This speech comprehension deficit co-occurred with fluent yet incomprehensible speech production ("paraphasias") due to replacements of speech sounds and entire words. In contrast to patients suffering from damage to Broca's area, however, these patients can produce grammatically intact sentences that make little sense. "*I was over in the other one, and then after they had been in the department, I was in this one*" is an example of the type of sentence produced by patient suffering from Wernicke's aphasia (Geschwind, 1970). These patients further seem to poorly realize that others have difficulty comprehending their speech. For an illustration of the anatomic locations of the Broca's and Wernicke's areas, as well as the arcuate fasciculus that connects these two areas, see **Figure 9-2** below.

There is a high degree of variability in the location of Broca's area

In a recent study, a substantial number of patients underwent mapping of the Broca's and Wernicke's areas based on pre-surgical electric stimulation. It was observed that the areas critical for speech production (tested with a naming task) and comprehension were more focally localized in individual patients than what has been believed based on previous studies on brain-lesioned patients. However, there was substantial inter-individual variability in the precise loci of the areas deemed critical for the language functions that were tested (Ojemann et al., 2008). This inter-individual variability in localization of brain functions is one of the major challenges also in neuroimaging studies where activity patterns during specific tasks are inspected at the group level.

The Wernicke-Geschwind model of speech

Inspired by the findings of Carl Wernicke and those of Paul Broca, another neuroscientist Norman Geschwind later developed an influential model of speech perception and production (for a detailed account of specific findings and predictions, see (Geschwind, 1970)). This model has become widely known as the Wernicke-Geschwind model. In short, the model assumes that when one listens to spoken words, the sound input traverses through the primary auditory cortex to the Wernicke's area where understanding of what was heard takes place.

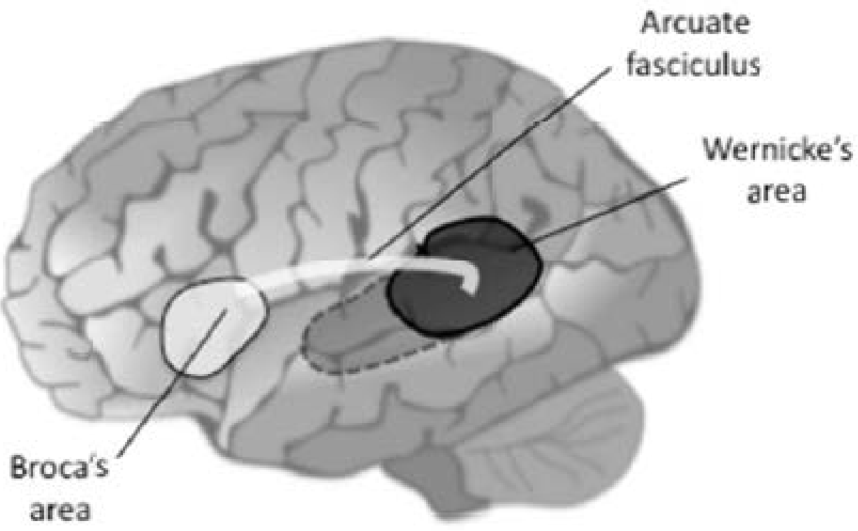


Figure 9-2: The approximate locations of the Broca's area, the Wernicke's area, and Arcuate fasciculus, the white-matter tract that anatomically connects the Broca's and Wernicke's areas. Note that here the area encompassing the junction between posterior temporal lobe and inferior parietal lobe, which is customarily referred to as Wernicke's area, is depicted with a darker shade. The lighter shade with the dotted outline extending anterior from this area indicates the area that is often left out when describing the Wernicke's area, however, this was included in the original descriptions of Carl Wernicke and also modern neuroimaging studies have supported the role of more anterior lateral temporal lobe areas in processing of speech stimuli (DeWitt and Rauschecker, 2012).

In the model, when speaking, the Wernicke's area communicates information *via* the arcuate fasciculus to the Broca's area that then converts the inputs into speech motor commands that trigger *via* the motor system the relevant sequence of articulations. Reading, on the other hand, involves passing of information from the visual cortex to Wernicke's area. Cutting of connections between the Wernicke's area and certain other brain areas was predicted (and partially also documented) to lead to specific language disturbances such as inability to read (Geschwind, 1970). Another symptom that patients with damage to a tract connecting the Broca's and Wernicke's areas, called arcuate fasciculus, suffer from is inability to repeat what they hear. This specific problem is called conduction aphasia.

Today, it is known that the Wernicke-Geschwind model is too simple, but it serves as a good example of how a theoretical model can be used to form testable hypotheses in cognitive neuroscience. Modern neuroimaging studies have produced a lot of evidence contrasting the Wernicke-Geschwind model. For instance, as will be described later in this chapter in more detail, it appears that decoding of speech based on the auditory input alone takes place in the anterior/ventral parts of the temporal lobe and not the posterior aspect of the temporal lobe that the classical Wernicke's area encompasses. Furthermore, recent neuroimaging have highlighted the role of the speech motor system in speech perception (see below).

Left hemisphere language dominance

As a careful reader has probably noted already, the findings of Broca and Wernicke (and also those of subsequent researchers) have documented language disturbances mainly following damage to the left hemisphere of the brain. Indeed, differences in the roles played by the left and right hemisphere is a robust finding in patient studies. In the vast majority of patients (> 90%) the left hemisphere is dominant for language. This figure depends on handedness: in 50% of left-handed the right hemisphere is the speech dominant hemisphere.

The speech dominance of the left hemisphere is manifested in left-hemisphere damage leading to severe speech and language deficits, both in terms of speech/language comprehension and production, more often than right hemisphere lesions. While there are studies that have noted more pronounced language lateralization in males than females (*e.g.*, males putatively exhibiting more severe disruptions of language functions following unilateral brain damage than females), careful analyses of the existing literature suggest that there are no robust sex differences in lateralization of language functions (Wallentin, 2009).

Findings of left-hemisphere language dominance obtained in brain-damaged patients are supported by observations in the so-called Wada test. In this test, a sedative barbiturate is administered through internal carotid arteries to one of the cerebral hemisphere at a time. This leads to inactivation of the targeted hemisphere that allows testing for language abilities of each hemisphere separately (Wada and Rasmussen, 1960). In these studies, it has been noted that inactivation of the left hemisphere results in loss of speech and language comprehension far more frequently than inactivation of the right hemisphere.

Neuroimaging studies using a variety of language tasks have provided more mixed results, with the between-hemisphere differences in the extent of activity during language tasks sometimes being surprisingly meager. However, there are findings suggesting that when a combination of different types of language tasks are utilized, neuroimaging results do provide similar language-lateralization results as could be predicted based on findings from patient and Wada test studies (Ramsey et al., 2001). It has to be noted that while the left hemisphere is more significant for language functions than the right in the vast majority of individuals, the right hemisphere does have at least rudimentary language capabilities and participates in language processing. Furthermore, the right hemisphere is more critical for certain aspects of language than the left hemisphere, such as emotional prosody/gesturing (Ross, 1979) and singing (Geschwind, 1970).

Language can be disturbed in several different ways

Besides the classic Broca's and Wernicke's aphasias, there are other ways in which speech and language functions can be disrupted following focal brain damage. Pure word deafness refers to a state in which the patient has relatively preserved auditory discrimination abilities but cannot comprehend words / speech sounds, and this problem can be highly specific (Saffran et al., 1976). Neuroanatomically, pure word deafness has been associated with damage to the middle temporal gyrus, but it can be also caused by damage to connectivity between brain areas (Binder et al., 2000).

Anomia refers to a condition where the patient is not able to recall words and names; unlike naming difficulties exhibited by patients with Broca's aphasia, there are not necessarily problems with grammar in anomia, and the patients often can describe what they mean in roundabout ways. Alexia means inability to read following brain damage, typically to the speech-dominant left hemisphere, and similar problems (of varying difficulty) are seen in developmental dyslexia. Agraphia, in turn, is the inability (or difficulty) to write that cannot be explained by general reduction in intelligence, paralysis or other secondary impairment.

Disruption of word fluency, a test where the subject is asked to produce within a given short time frame as many words as possible that begin with a specified letter (Borkowski et al., 1967), typically follows from damage to the frontal lobe and is manifested as impoverished speech (Henry and Crawford, 2004). Lack of words generated that begin with the designated letter, violations of the rule (occasionally generating words that begin with some other letter), and perseveration errors (repeating the same word) are each the type of problems manifested in this task in frontal-lobe patients. Supporting the brain lesion findings, healthy subjects were observed in neuroimaging studies to exhibit activity in left inferior frontal gyrus, anterior cingulate gyrus, and superior frontal sulcus activity during a word fluency task, as contrasted with word repetition and antonym generation tasks (Phelps et al., 1997).

Right-hemisphere language functions

While most of the language problems typically follow left-hemisphere injury, damage to the right hemisphere, especially to the frontal lobe, can result in lack of prosodia (Ross, 1979). This means that the patient speaks in a flat non-emotional voice, with emotional expression such as exclamations and emotional tone lacking. There are also curious observations relating to right hemisphere role in language, as some patients with severe Broca's aphasia, while unable to speak, can fluently communicate their thoughts by singing (Geschwind, 1970). Importantly, the broad range of specific language and speech deficits illustrates the multifaceted nature of language ability; instead of a single language function, there are several functions that are at least partially served by different brain regions and/or distinct networks of brain areas.

The N400 response is elicited by semantic incongruence

The invention and rapid development of non-invasive neuroimaging methodology, such as positron emission tomography, functional magnetic resonance imaging, electroencephalography, and magnetoencephalography (see Chapter 2), is increasingly making it possible to investigate the neural basis of speech and language functions in healthy volunteers by measuring brain activity patterns during various language tasks. By recording electroencephalogram epochs time-locked to onset of words in a sentence, Marta Kutas and Steven Hillyard observed that there was a specific response beginning at about 250 ms, and peaking at approximately 400 ms, from onset of a semantically incongruent word in the sentence (*e.g.*, "He took a sip from the transmitter") (Kutas and Hillyard, 1980). This response, which was named the N400 response, is shown in Figure **9-3**. It is currently believed that the N400 response is generated as the result of forming semantic predictions, based on the context and preceding words, in a network of brain areas involving middle temporal gyrus, anterior temporal lobe, angular gyrus, and inferior frontal gyrus (Lau et al., 2008).

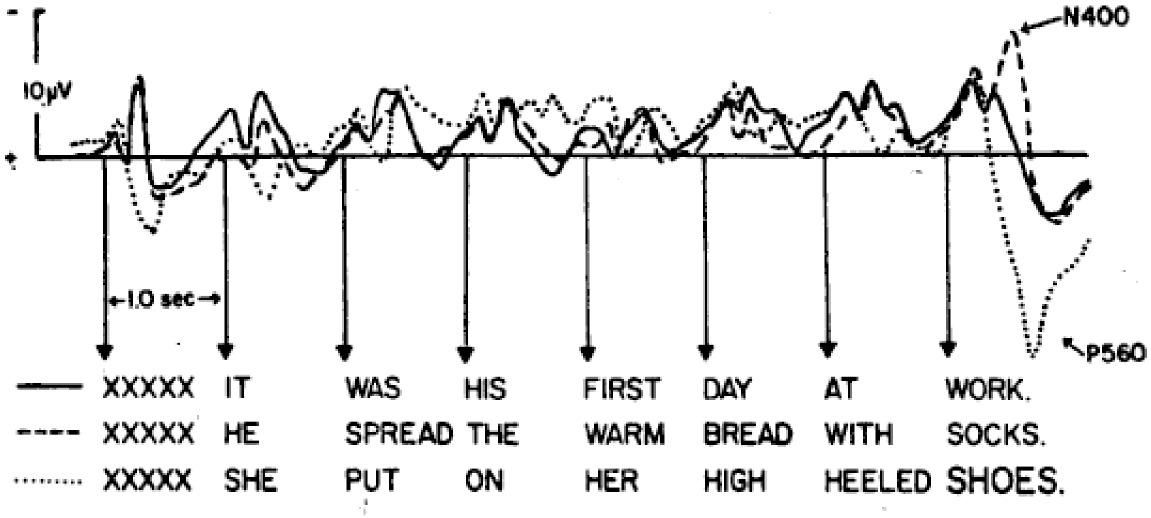


Figure 9-3. Across-trial averaged electroencephalogram epochs time-locked to onsets of words in a sentence disclose a specific response called N400 to a semantically incongruent word in a sentence (*e.g.*, "He spread the warm bread with socks" instead of "It was his first day at work"). When the seventh word was presented with larger font size, there was another response called the P560 response. Adapted from (Kutas and Hillyard, 1980).

Speech sounds activate anterior-lateral parts of the temporal lobe

One robust finding that seems to be rather consistent across neuroimaging studies is that, seemingly contradictory with previous patient studies, the relative sensitivity of the brain to non-verbal *vs*. verbal stimulation activates regions anterior and ventral to the primary auditory cortex on the lateral surface of the temporal lobe, as opposed to the temporal-parietal junction that has been traditionally considered the Wernicke's area (Binder et al., 2000). For a recent meta-analysis supporting these observations, see (DeWitt and Rauschecker, 2012). The results of this meta-analysis are presented in **Figure 9-4**. In fact, these findings match quite well what Wernicke originally described as the speech receptive area; it is subsequent literature that has referred temporal-parietal junction as Wernicke's area (the approximate area that has been left out is shown as the lighter-shaded region in **Figure 9-2** above). For an in-depth analysis of the original texts and how the concept Wernicke's area has evolved, see (Bogen and Bogen, 1976).

There are neuroimaging studies that have found speech-specific activations in the posterior temporal lobe areas. **Figure 9-5** shows one example of a study that has found posterior temporal lobe activity during speech perception (Möttönen et al., 2006). In this study, subjects were presented so-called sine-wave speech sounds; such sounds are generated by positioning time-varying sine waves at the center frequencies of three specific frequency bands of recorded speech sounds. These frequency bands, also known as formants, carry speech-specific information. When one is not informed that the sounds are in fact speech sounds, they sound like chirping noises, but as soon as one is informed that the sounds are speech sounds, they are perceived as speech. When brain hemodynamic activity was contrasted between informed and non-informed conditions (*i.e.*, when the subjects heard the sine-wave speech stimuli as speech *vs*. non-speech) the posterior temporal lobe areas were activated in the left hemisphere.

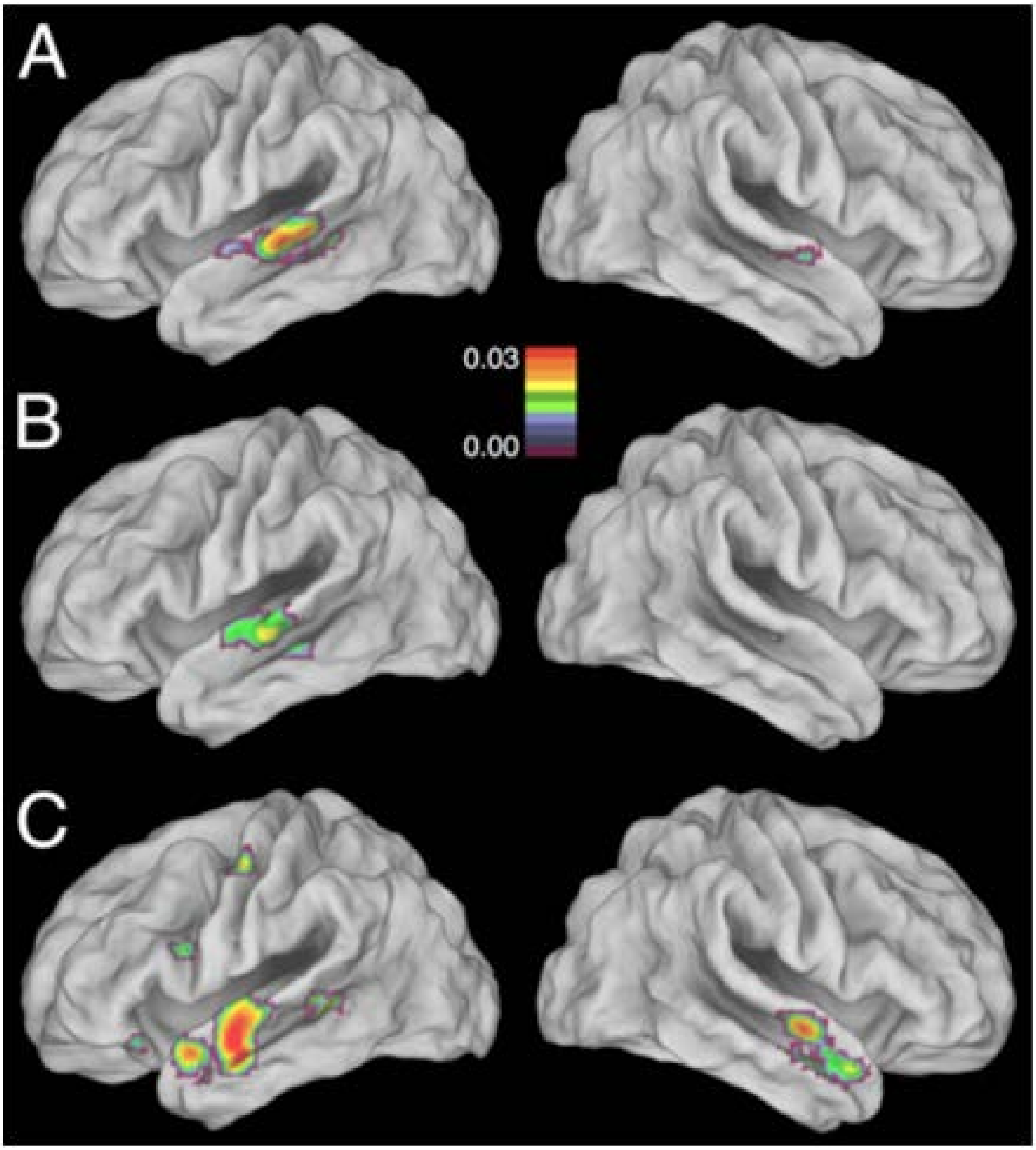


Figure 9-4: Shown are results from a recent meta-analysis of neuroimaging findings that show left-hemisphere dominance and progression of activations towards anterior/ventral direction along the temporal lobe during processing of A) phonemes, B) words, and C) phrases. The color code plots, on the cortical surface, regions estimated as being statistically significantly involved across studies. Adapted from (DeWitt and Rauschecker, 2012).

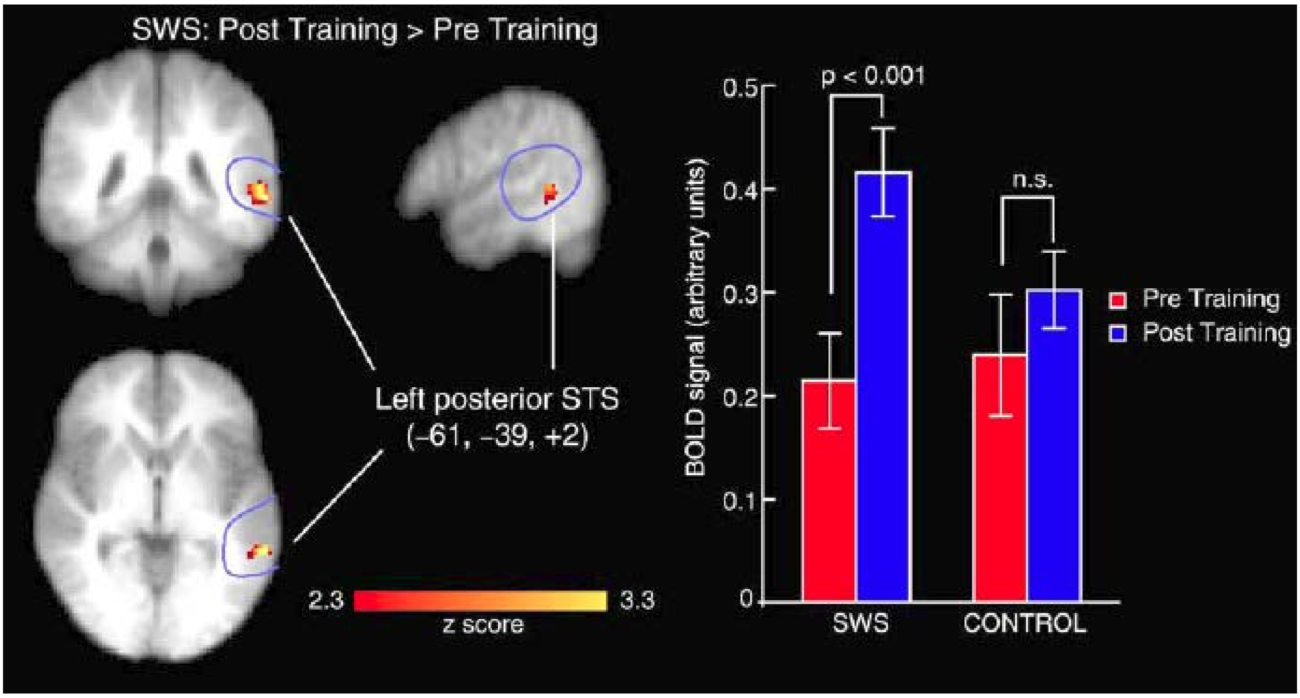


Figure 9-5. Activation of the left posterior temporal lobe in the vicinity of the classical Wernicke's area when sine-wave speech was perceived as speech vs. non-speech noise. See text for details. Adapted from (Möttönen et al., 2006).

It is increasingly recognized that posterior temporal / parietal activations during speech perception result from mapping of speech sounds to motor information of how speech sounds are articulated, in parallel to analysis of speech sounds based on auditory features that takes place anterior/ventral to the primary auditory cortex. This so-called motor theory of speech perception is a relatively old concept (Liberman et al., 1967) and areas posterior to the primary auditory cortex, inferior parietal lobe and motor areas of the frontal cortex seem to be essential for processing of speech based on mapping of sounds to articulatory code (Hickok and Poeppel, 2004, Rauschecker and Scott, 2009, Jaaskelainen, 2010). This theory and related empirical findings are briefly presented in the following.

9.4 The motor theory of speech perception

When one inspects the acoustic properties of speech sounds that are revealed by spectrograms (*i.e.*, plots of acoustic power as a function of sound frequency and time), it soon becomes evident that the constituent speech sounds, called phonemes, are not invariant but rather their spectrotemporal composition depends on the phonetic context in which they are pronounced. This phenomenon is known as co-articulation, and it is illustrated in **Figure 9-6** below where /s/ is acoustically different in "seat" *vs*. "suit". Specifically, since the /s/ does not require use of lips, lip movements related to pronunciation of adjacent sounds (such as rounding of lips during /u/ in suit) can begin during the /s/. Another good example of the ambiguity created by co-articulation is that /p/ in "port" is rather different from /p/ in "sport", yet the phonetic identity of the /p/ is retained.

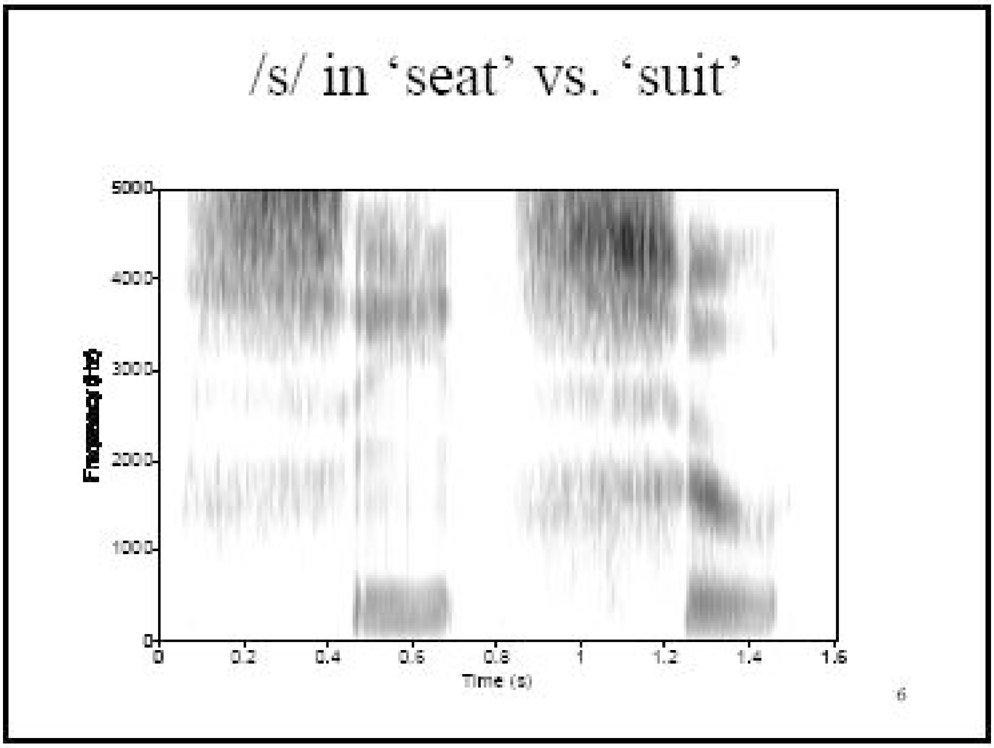


Figure 9-6. Shown is an example of the co-articulation phenomenon where speech sounds differ based on the phonetic context where they are articulated in. Here the spectrogram of two speech sounds "seat" and "suit" that both begin with /s/, show how the /s/ in these two speech sounds differs acoustically. Spectrogram plots fluctuations of energy as a function of time over frequencies that contain speech information, with the darker shade here representing higher acoustic power (adapted from http://cueflash.com/cardimages/questions/thumbnails/0/9/4890745.jpg).

Observations such as this has led speech scientists to speculate that knowledge of the motor/articulatory sequences could be what the auditory input is mapped onto when one perceives speech (Liberman et al., 1967). The speech motor theory, and the seemingly conflicting findings of anterior/ventral *vs*. posterior temporal lobe activations during speech perception tasks in neuroimaging studies, has inspired recent empirical and theoretical work to propose dual-stream models of speech perception (Hickok and Poeppel, 2004, Basirat et al., 2008, Rauschecker and Scott, 2009, Jaaskelainen, 2010). In these models, speech is processed based on purely acoustic analysis in areas anterior/ventral to the primary auditory cortex, whereas acoustic speech features are mapped onto motor schemata in the dorsal processing pathway that involves posterior temporal lobe and inferior parietal areas, as well as motor areas in the frontal cortex. Speech perception then is presumed to take place as a result of interactive resonance within and across these two streams.

Integration of auditory visual and motor speech cues during development

It has also been argued that during development an infant receives temporally coincident visual, auditory, and motor inputs during speech perception (as small babies tend to attempt to imitate the speech of their caretakers while also tending to intently watch the face of the caretaker). Given that these inputs converge in the posterior temporal areas rather than anterior/ventral areas due to anatomical inter-connectivity of the auditory, visual, and motor systems, it is possible that neuronal receptive fields in areas posterior to the primary auditory cortex are shaped to detect different aspects of the acoustic speech inputs (*i.e.*, those that correlate with motor and visual cues) than the anterior stream that only receives auditory inputs (Jaaskelainen, 2010).

Indeed, it is well known that visual speech (the position of lips and tongue) can give complementary information that helps disambiguate auditory speech especially in noisy conditions (Sumby and Pollack, 1954), and that incongruent visual information can produce illusory speech percepts (McGurk and MacDonald, 1976). For certain speech sounds, the position of the tongue and lip forms convey less ambiguous information than the speech sounds themselves. The origin of this effect is likely visual-auditory interactions already in the primary auditory cortex (Sams et al., 1991, Calvert et al., 1997) and the speech motor system has been suggested to play a central role in audio-visual speech integration as well.

9.5 Distributed representations underlie concepts

Neuroimaging studies mapping distributed patterns of brain activity that correspond to concept representations have recently yielded further support for the notion that networks consisting of multiple brain areas, and/or processing streams, underlie human language. **Figure 9-7** shows an example of findings where patterns of activity corresponding to concepts (*e.g.*, dwellings *vs*. tools) consistently replicate across subjects and across languages, thus revealing multifocal concept representations in the brain. These findings demonstrate that language representations in the brain are not restricted to a single cortical "language area", but rather that it is the activity in several brain locations that underlies language representations.

What is notable about the concept representations is that the concepts of tools, for instance, involve activation of brain areas that are utilized during motor behavior (*i.e.*, tool use). This suggests that concept representations are at least to some extent embodied. Furthermore, the fact that semantically similar concepts (such as dwellings and tools) activate similar patterns of distributed areas potentially explains the results from behavioral studies (described earlier in this chapter) where semantically similar words primed the processing of one another. It seems that semantically similar concepts are adjacent also in the brain space (perhaps due in part to the ways that they are embodied) and thus spreading of activity could lower the neural threshold for processing a subsequently presented semantically related concept.

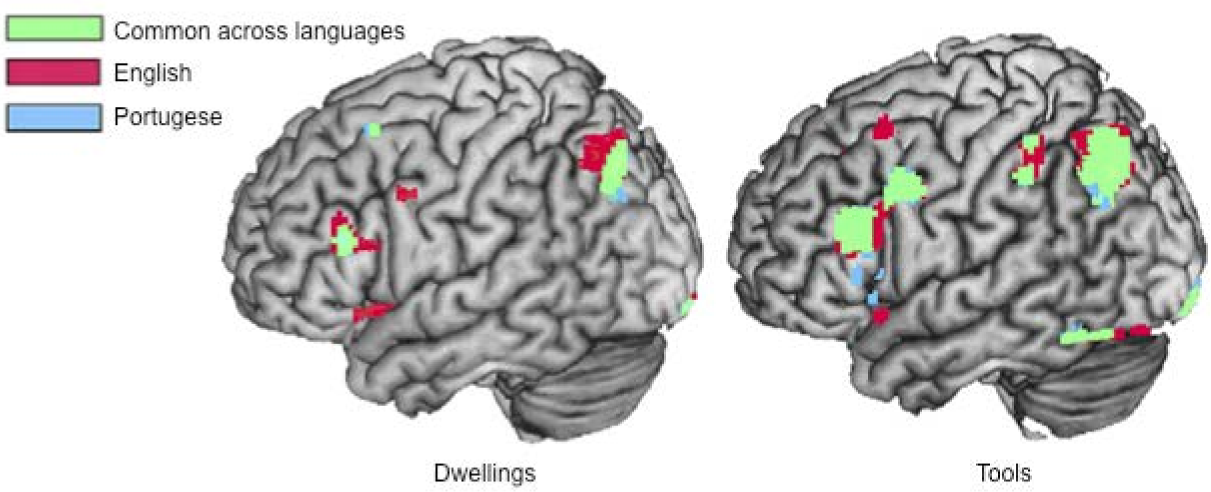


Figure 9-7: Neuroimaging of distributed activity patterns underlying concept representations. Shown here are brain regions consistently activated across subjects when thinking about the properties of tools and dwellings in two distinct languages, English and Portuguese. The common regions are plotted with light green color. Adapted from (Buchweitz et al., 2012).

What also makes these findings interesting is that the distributed brain activity patterns can be successfully utilized in subsequent crude forms of "mind reading". Specifically, after having first recorded brain activity during presentation of concepts in different (*e.g.*, word and pictorial) forms using functional magnetic resonance imaging, machine-learning algorithms can be utilized to find common distributed sets of brain areas that represent the concepts. In subsequent recordings, the distributed activity patterns can be utilized to correctly identify (*i.e.*, "mind read") concepts that the subjects are engaged in thinking about (Just et al., 2010). Of course, such mind reading methods are presently still quite limited, but they are good indications of the thrilling possibilities that development of non-invasive neuroimaging techniques are making possible in cognitive neuroscience.

9.6 Concluding remarks

It can be said that language is a unique human ability that has made it possible for phenomena such as abstract thinking, shared concepts, intricate social interactions and relationships, sophisticated cultures, and societies to emerge. Early observations of specific speech and language deficits in patients with focal brain lesions paved way for the study of the neural basis of language. Modern neuroimaging experiments have revealed multiple underlying brain areas and parallel streams for processing of speech sounds. It has become evident that knowledge of speech motor sequences plays a role in speech perception in addition to speech production. Finally, it has been shown recently that concepts are represented in the brain as distributed multifocal activity patterns. Importantly, language supports other aspects of cognition, such as formulating goals that direct one's behavior and solving problems that one encounters when pursuing such goals. The cognitive functions that support goal-directed behavior are collectively called executive functions. Executive functions and the underlying neural mechanisms will be introduced in the next chapter.

9.7 References

BASIRAT, A., SATO, M., SCHWARTZ, J. L., KAHANE, P. & LACHAUX, J. P. 2008. Parieto-frontal gamma band activity during the perceptual emergence of speech forms. *Neuroimage,* 42**,** 404-413.

BINDER, J. R., FROST, J. A., HAMMEKE, T. A., BELLGOWAN, P. S., SPRINGER, J. A., KAUFMAN, J. N. & POSSING, E. T. 2000. Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex,* 10**,** 512-528.

BOGEN, J. E. & BOGEN, G. M. 1976. Wernicke's region-where is it? *Annals of the New York Academy of Sciences,* 280**,** 834-843.

BORKOWSKI, J. G., BENTON, A. L. & SPREEN, O. 1967. Word fluency and brain damage. *Neuropsychologia,* 5**,** 135-140.

BROCA, P. P. 1861. Perte de la parole, ramollissement chronique et destruction partielle du lobe antérieur gauche du cerveau. *Bulletin de la Société Anthropologique,* 2**,** 235-238.

BUCHWEITZ, A., SHINKAREVA, S. V., MASON, R. A., MITCHELL, T. M. & JUST, M. A. 2012. Identifying bilingual semantic neural representations across languages. *Brain and Language*, 120, 282-289.

CALVERT, G. A., BULLMORE, E. T., BRAMMER, M. J., CAMPBELL, R., WILLIAMS, S. C., MCGUIRE, P. K., WOODRUFF, P. W., IVERSEN, S. D. & DAVID, A. S. 1997. Activation of auditory cortex during silent lipreading. *Science***,** 593-596.

COLLINS, A. M. & LOFTUS, E. F. 1975. A spreading-activation theory of semantic processing. *Psychological Review,* 82**,** 407-428.

DEWITT, I. & RAUSCHECKER, J. P. 2012. Phoneme and word recognition in the auditory ventral stream. *Proc Natl Acad Sci USA*, 109, E505-E514

FORSTER, K. I. & CHAMBERS, S. M. 1973. Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior,* 12**,** 627-635.

GARDNER, R. A. & GARDNER, B. T. 1969. Teaching sign language to a Chimpanzee. *Science,* 165**,** 664-672.

GESCHWIND, N. 1970. The organization of language and the brain. *Science,* 170**,** 940-944.

HENRY, J. D. & CRAWFORD, J. R. 2004. A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology,* 18**,** 284-295.

HICKOK, G. & POEPPEL, D. 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition,* 92**,** 67-99.

JAASKELAINEN, I. P. 2010. The role of speech production system in audiovisual speech perception. *Open Neuroimaging Journal,* 4**,** 30-36.

JUST, M. A., CHERKASSKY, V. L., ARYAL, S. & MITCHELL, T. M. 2010. A neurosemantic theory of concrete noun representation based on the underlying brain codes. *PLoS ONE,* 5**,** e8622.

KUTAS, M. & HILLYARD, S. A. 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science,* 207**,** 203-205.

LAU, E. F., PHILLIPS, C. & POEPPEL, D. 2008. A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience,* 9**,** 920-933.

LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P. & STUDDERT-KENNEDY, M. 1967. Perception of the speech code. *Psychological Review,* 74**,** 431-461.

MCGURK, H. & MACDONALD, J. 1976. Hearing lips and seeing voices. *Nature,* 264**,** 746-748.

MÖTTÖNEN, R., GALVERT, G. A., JAASKELAINEN, I. P., MATTHEWS, P. M., THESEN, T., TUOMAINEN, J. & SAMS, M. 2006. Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. *Neuroimage,* 30**,** 563-569.

NEELY, J. H. 1977. Semantic priming and retrieval from lexical memory: roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General,* 106**,** 226-254.

OJEMANN, G. A., OJEMANN, J., LETTICH, E. & BERGER, M. 2008. Cortical language localization in left, dominant hemisphere. An electrical stimulation mapping investigation in 117 patients. *J Neurosurg,* 108**,** 411-421.

PAGEL, M., ATKINSON, Q. D. & MEADE, A. 2007. Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature,* 449**,** 717-720.

PHELPS, E. A., HYDER, F., BLAMIRE, A. M. & SHULMAN, R. G. 1997. FMRI of the prefrontal cortex during over verbal fluency. *NeuroReport,* 8**,** 561-565.

RAMSEY, N. F., SOMMER, I. E. C., RUTTEN, G. J. & KAHN, R. S. 2001. Combined analysis of language tasks in fMRI improves assessment of hemispheric dominance for language functions in individual subjects. *Neuroimage,* 13**,** 719-733.

RAUSCHECKER, J. P. & SCOTT, S. K. 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci,* 12**,** 718-724.

ROSS, E. D. 1979. Dominant language functions of the right hemisphere? Prosody and emotional gesturing. *Archives of Neurology,* 36**,** 144-148.

SAFFRAN, E. M., MARIN, O. S. & YENI-KOMSHIAN, G. H. 1976. An analysis of speech perception in word deafness. *Brain and Language,* 3**,** 209-228.

SAMS, M., AULANKO, R., HAMALAINEN, M., HARI, R., LOUNASMAA, O. V., LU, S. T. & SIMOLA, J. 1991. Seeing speech: visual information from lip movements modifies activity in the human auditory cortex. *Neurosci Lett,* 127**,** 141-145.

SUMBY, W. H. & POLLACK, I. 1954. Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America,* 26**,** 212-215.

WADA, J. & RASMUSSEN, T. 1960. Intracarotid injection of sodium amytal for the lateralization of cerebral speech dominance: experimental and clinical observations. *Journal of Neurosurgery,* 17**,** 266-282.

WALLENTIN, M. 2009. Putative sex differences in verbal abilities and language cortex: A critical review. *Brain and Language,* 108**,** 175-183.

WERNICKE, C. 1874. *Der aphasische symptomencomplex: eine psychologische studie auf anatomischer basis,* Breslau, Cohn & Weigert.

10 Executive functions

Executive functions refer to a range of higher-order cognitive functions that enable organized and goal-directed behavior. Problems manifested by two exemplary brain damaged patients help illustrate the role executive functions play in cognition and everyday life. The first example is patient A, who suffers brain damage in a traffic accident. In the hospital he seems to recover well. He is able to perform relatively well in standard intelligence quotient tests as well as in other tests that are used in clinical neuropsyhological examination, such as tests for memory and attention. Surprising problems emerge, however, during the weeks following release from hospital. Relatives soon begin to complain that the patient's behavior seems impulsive and erratic, selling his luxurious home for few thousands, forgetting to pay items that he takes from the store, and ending up in quarrels with his friends and neighbors.

Another exemplary patient, here called patient B, suffers a stroke. After initial recovery he too shows surprisingly intact performance in standard clinical neuropsychological tests. It does not, however, take long before his relatives start complaining about what they describe as apathy and lack of initiative exhibited by the patient at home. He fails to initiate and carry through any behaviors on his own. However, if instructed and supervised through a sequence of behaviors such as preparing and eating breakfast he carries out the tasks without major problems. What these exemplary patient cases have in common is that both exhibit problems in executive functions that often follow from damage to the frontal lobes of the brain, especially to prefrontal cortical areas. Findings pertaining to this will be described in the following.

10.1 Prefrontal cortex is central for executive functions

19th Century railroad construction worker named Phineas Gage is perhaps the very first case in which deficits in executive functions have been documented following damage to the anterior part of the brain. As described by John Harlow, the physician who treated Gage after the accident, on the day of the accident Gage "… was engaged in charging a hole, preparatory to blasting. He had turned in the powder, and was in the act of tamping it slightly before pouring on the sand. He had struck the powder, and while about to strike it again, turned his head to look after his men (who were working within a few feet of him), when the tamping iron came in contact with the rock, and the powder exploded, driving the iron against the left side of his face … taking a direction upward and backward toward the median line, it … entered the cranium, passing through the anterior left lobe of the cerebrum, and made its exit in the median line … fracturing the parietal and frontal bones extensively, breaking considerable portions of the brain, and protruding the globe of the left eye from its socket, by nearly one half its diameter" (Harlow, 1848).

Miraculously, Gage survived the blast, but the man who had prior to the accident been the most responsible and determined (in Harlow's words, based on the accounts of people who knew him, Gage was "*of temperate habits, and possessed considerable energy of character*"), became impulsive, unable to set and pursue goals, hold employment etc. His general intelligence and memory functions, however, remained seemingly unaffected by the damage. The location and extent of the lesion was subsequently reconstructed in 1994 based on how the skull of Gage had been damaged (Damasio et al., 1994). This is shown in **Figure 10-1** below.

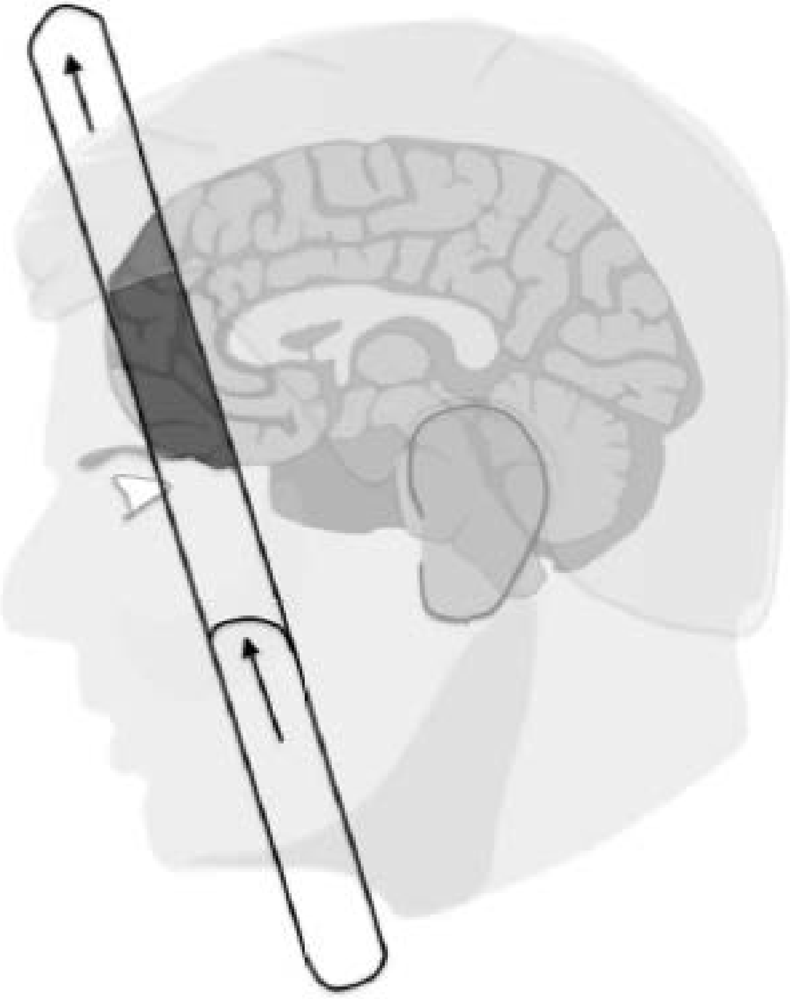


Figure 10-1. An illustration of the damage caused by blasting of the tampering rod through the anterior parts of Phineas Gage's brain in a railroad construction accident in 1848, drawn based on results presented in (Damasio et al., 1994).

The prefrontal cortex is a heterogeneous area

In subsequent studies, the frontal lobes have been well documented to be integral for a variety of executive functions, including social cognition (that will be introduced in Chapter 12). In fact, the prefrontal cortex is far from being an anatomically and functionally uniform area (see Figure 10-2). It is, however, important to keep in mind that the prefrontal cortex cannot alone account for the executive functions; the role of the prefrontal cortex is more that of a central hub in a network of brain areas that underlie executive functions. The connectivity pattern of prefrontal cortex is ideally suited for this task, as there are reciprocal connections to most other brain areas. Thus it is not surprising that damage to prefrontal cortex, or the connections that link this area to the rest of brain, easily result in failure of executive functions.

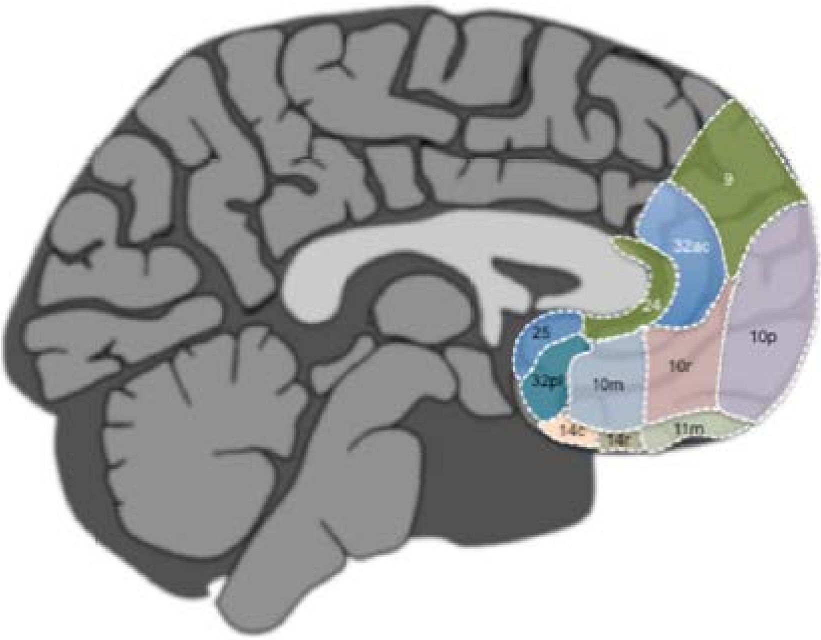


Figure 10-2. Anatomic subregions of the human anterior-ventral prefrontal cortex based on cytoarchitectonic differences. This fine parcellation of the anatomic subregions is also manifested in what specific roles each of the areas play in cognitive functions. Drawn based on results presented in (Öngur et al., 2003).

Notably, outside the scope of problems encountered in persons with grave brain damage such as that suffered by Phineas Gage, there are differences between healthy individuals in how well one's behavior is overall organized, how flexibly one can alter plans when obstacles emerge, and partly as a result such differences also how much a given individual can accomplish in life. The constituents of executive functions, such as how flexibly one can take negative feedback into account and how persistently one can pursue life goals, and the underlying neural mechanisms, will be introduced in the following.

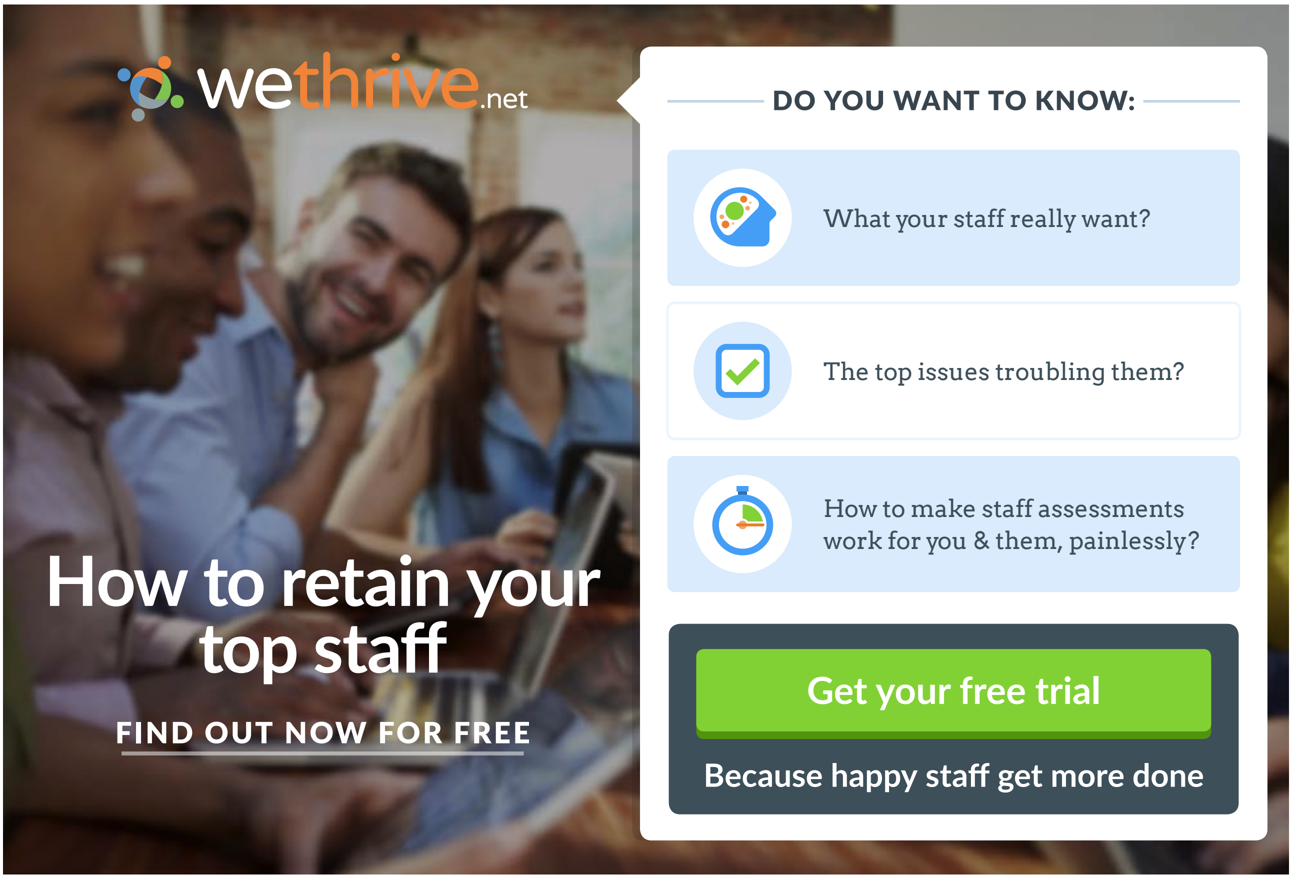
10.2 The constituents of executive functions

There are multiple executive functions, for instance, setting goals and plans, as well as appropriate sequencing of actions to attain those goals, is a very important aspect of executive functions. In fact, executive functions are often even referred to with the term goal-directed behavior. Another important ability that is at the core of executive functions is the ability to suppress (or inhibit) competing goals from disturbing or diverting the individual from pursuing the concurrently relevant goal. As you might recall from Chapters 6 and 7, ability to inhibit processing of irrelevant stimuli was described as an important facet of selective attention, and that the ability to select and manipulate relevant material was described as being a central aspect of working memory. Indeed, selective attention and working memory are often considered as important components of executive functions. Internal speech and ability to self-reflect are also considered as executive functions.

At a hierarchically low level, motor execution, guiding and sequencing of movements can be considered part of executive functions. Related to this, it seems that the areas planning motor sequences also make it possible to mentally plan sequences of actions. The schemas that were introduced in Chapter 7 are also a central concept in executive functions, and emotions that will be described in more detail in Chapter 11 are fueling motivation that is crucial for setting of goals and perseverance in attaining them. In fact, it is a good exercise to try to think of any important life goals that would not have something to do with emotions – if one thinks about it carefully, there probably are none. We are intrinsically bound and guided by emotions in most aspects of our behavior. Finally, a topic of increasingly intensive research in cognitive neuroscience, decision making, is an intimate aspect of executive functions. In the following, these constituents of executive functions will be introduced.

Setting and keeping goals across different timescales

As can be seen in **Figure 10-3** below, in order to orderly attain any major goal in life, one needs to set a hierarchy of goals, with the hierarchically highest goals spanning over much greater timescales than the minute subgoals that relate to, for instance, daily survival such as buying food from the store and eating a meal. In the example outlined in the figure, the desire to become a medical doctor is a representative of a longer-term goal of a college student. The subgoals that then follow from this overall goal are finding out about admission criteria to various medical schools, sending out applications, going to possible entrance examinations, and once admitted to a medical school, keeping focus on studies.



Furthermore, in order to become a skilled clinician, one needs to build a career as an M.D., and the first step in that process is applying for the first job as an M.D. Importantly, the ability to sequence subgoals in a coherent fashion is something that patients with prefrontal cortical damage often have difficulties in accomplishing. While they might be able to tell all the necessary subtasks that are needed in, for instance, preparing a meal, they fail to sequence them appropriately and/or suffer from intrusion of competing impulses (Lezak et al., 2004) (see also below the description of the so-called multiple errands test).

Of course, there are multiple motivational factors that help sustain the hierarchically higher-order goals. These naturally vary between individuals. For some, the motivation to become an M.D. might stem primarily from a desire to be able to help sick children, for others, factors such as wishing to get an esteemed profession might also play a role. Naturally, during the course of studies the motivational factors might also change, for instance, someone initially motivated by visions of being in an esteemed profession might be moved by a specific group of patients, such as when seeing the plight of patients suffering from cancer, and become intrinsically motivated to help.

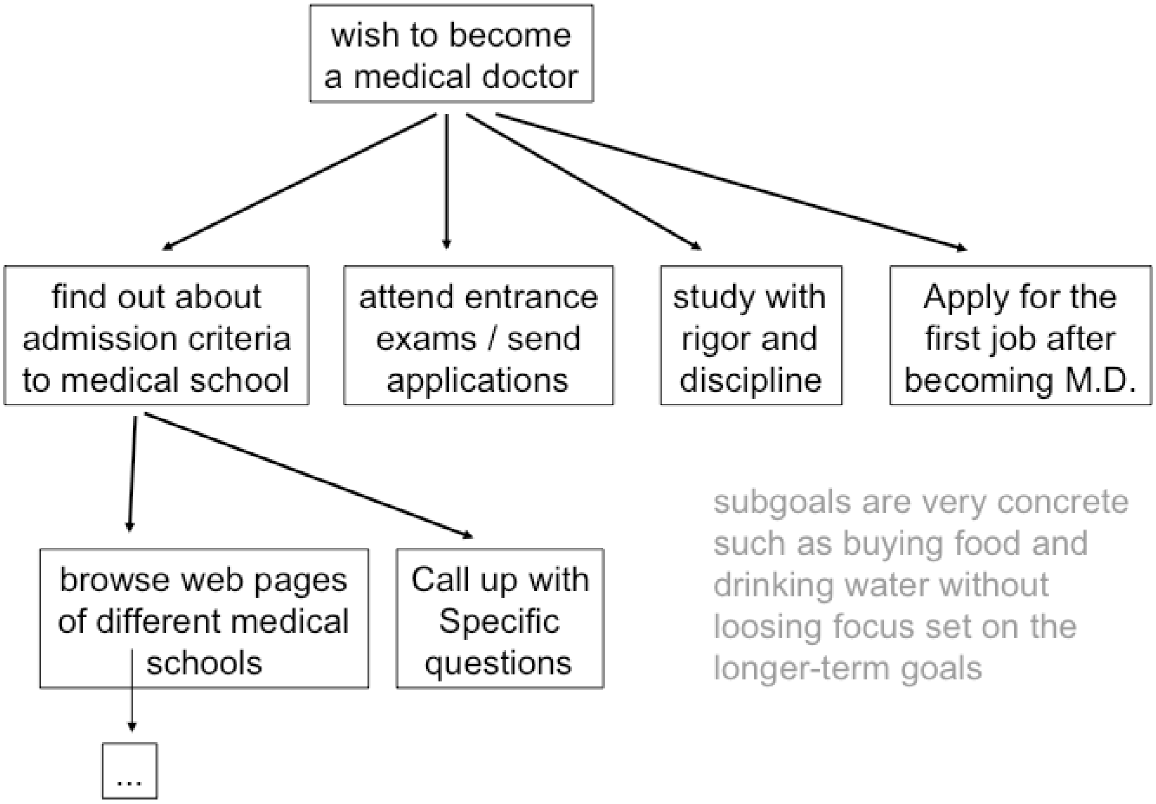


Figure 10-3. Shown is an example of hierarchy and sequencing of goals that are used to guide daily behaviors. Hierarchically higher-order goals are listed at the top and the sequential subgoals that serve the higher-order goals are listed below.

Inhibition of competing/interfering impulses and goals

The ability to inhibit interfering impulses and goals is another highly important aspect of executive functions. In healthy subjects, response inhibition has been empirically studied using reaction time tasks where a subject is to press a button to one set of stimuli and inhibit responding to another set of stimuli (these are called Go/NoGo tasks). Another behavioral test is the stop-signal task where a stop signal is delivered in a reaction time task; the duration of the stopping process is then taken as the measure of inhibitory control. Deficiency in the ability to inhibit competing responses, which is often caused by frontal lobe damage, especially to inferior aspects of the right frontal lobe (Aron et al., 2004), is manifested in daily life as inability to focus on concurrent tasks and in loose associations. Behavior is strongly stimulus-driven in these patients, as external stimuli (as well as internal thought processes) easily take over the focus of the patient and interrupts the ongoing behavior.

Failure to inhibit competing impulses/goals also contributes to difficulties in proper sequencing of tasks to attain any longer-term goals. For instance, when preparing a meal, the patient might get a competing impulse to watch television that he/she cannot suppress, followed by some other impulse, thus resulting in a failure to finish the dinner preparations. In children, there is a developmental disorder called attention-deficit hyperactivity disorder (ADHD) that is characterized by inability to inhibit interfering stimuli (Schachar et al., 1995). Further, in neuropsychological examination, ADHD patients overall resemble those with frontal lobe lesions (Shue and Douglas, 1992). For ADHD children it is, for instance, quite essential that the learning environment at school contain as few disruptive stimuli as possible, helping them to maintain focus of attention.

Ability to take feedback flexibly into account and adjust behavior

There are often situations where something comes up and it is no longer feasible to carry on with the current goal. Deciding to take an alternate route to work due to an unexpected traffic jam (*e.g.*, due to an accident blocking the road ahead) is a commonplace example of this. A healthy person can quickly figure out an alternative driving route, take a turn to a side street, and manage to arrive (more or less) timely at work. This involves abandoning the original plan and adopting a new route around the traffic accident, however, the overall higher-order goal of getting to work in time does not change.

Life goals can also be subject to the type of feedback that calls for adjustment of one's goals. For instance, if one is determined to get his/her M.D. degree from one of the most esteemed medical schools, fails to get admitted there, but does get selected to some other medical school, it might be feasible to change plans and commence studies elsewhere, rather than wasting time preparing to reapply with a substantial delay. It is good to note that radical changes in life goals can cause stress even if they are brought about by positive events such as winning tens of millions in a lottery.

Perseveration: an inability to adjust one's behavior

Some patients with prefrontal cortical damage exhibit problems in their ability to adjust their plans given negative feedback, but rather suffer from pathological persistence with a selected mode of action. This specific problem is called perseveration, for instance, a patient might break down a door that is locked instead of seeking an alternate route (note that with perseverance, on the other hand, one typically denotes appropriately determined and resilient pursuing of important goals). Positron emission tomography studies on closed head injury patients have suggested that damage to circuitry involving dorsolateral prefrontal cortex and striatum gives rise to perseveration problems (Lombardi et al., 1999).

The motor system supports executive functions

At the hierarchically lowest level of action execution are individual motor acts (*e.g.*, moving a limb, reaching for an object, and speaking). The motor system of the brain is, however, not limited to carrying out simple motor actions. Rather, as described in Chapter 5 there is a posterior-to-anterior hierarchy of brain areas in the frontal lobes responsible for increasingly complex motor sequences; this hierarchically organized cortical motor system has been observed to play a vital role in a number of cognitive functions such as speech perception. Executive functions are not an exception, as motor imagery and planning of actions (without necessarily executing them) are supported by the motor system (Jeannerod, 1995). Thus, planning how to accomplish specific goals is accomplished partly by the same brain regions that implement the action sequences.

An intricate network of brain areas consisting of prefrontal cortex, basal ganglia, and thalamus seems to be responsible for choosing between alternative actions and inhibiting competing motor plans and actions. Patients suffering from Parkinson's disease, who have lost the vast majority of dopaminergic inputs from substantia nigra to striatum, suffer from inability to initiate movements, but also from cognitive deficits. In particular, patients with Parkinson's disease suffer from cognitive inflexibility.

Specifically, Parkinson's disease patients have problems in flexibly shifting from one "cognitive set" to another (Gotham et al., 1988). The term cognitive set refers in cognitive neuroscience to a mode of action that includes the concurrent goal and the active schemas that guide behavior towards the goal. Recent neuroimaging evidence suggest that the deficiency in cognitive set shifting in Parkinson's patients is due to deficits in neural communication between the prefrontal cortex and striatum that depends on intact dopamine transmission in the brain (Monchi et al., 2004). The nuclei of basal ganglia and their role in action/cognitive set selection will be described next.

Basal ganglia: selection of actions and cognitive sets

Basal ganglia consist of a number of interconnected nuclei that are shown in **Figure 10-4**. One of these nuclei, substantia nigra (and specifically the pars compacta part of substantia nigra), is specifically implicated in Parkinson's disease. The role of substantia nigra pars compacta in selection of actions and cognitive sets is due to the dopaminergic connections that ascend from this region to striatum (hence named the "nigrostriatal" pathway) exerting excitatory influence on the so-called direct pathway, and inhibitory influence on the so-called indirect pathway.

The direct pathway is generally believed to be responsible for enhancing selection of a desired action or cognitive set, and the indirect pathway for suppressing irrelevant / disruptive actions/sets. The influence of the nigro-striatal excitatory inputs on the direct pathway are based on the dopaminergic neuronal tracts that ascend from substantia nigra pars compacta and excite striatum. Striatum then inhibits internal globus pallidus, reducing the tonic inhibition that internal globus pallidus exerts on ventral lateral and ventral anterior nuclei of the thalamus. This results in the ventral lateral and anterior lateral nuclei of the thalamus sending excitatory inputs to cortex.

As can be seen in **Figure 10-4**, the net effect on the internal segment of globus pallidus *via* the indirect pathway is also one that reduces inhibition at the level of thalamus. As can also be seen in this case, the effect of dopamine (excitatory *vs*. inhibitory) depends on the type of postsynaptic receptors, and not on the neurotransmitter substance *per se* (this principle was described in Chapter 4). One of the largest challenges in drug development is to identify molecules that would specifically bind to only a certain class of receptors. The increased specificity provides the advantage of more precisely producing the desired effect, whilst avoiding unwanted side effects that are caused when the drug binds to other receptors besides the ones that are targeted by the drug.

A state in many ways converse to the symptoms of Parkisons's disease is seen during acute psychotic symptoms in a schizophrenia patient. It is hypothesized that especially the so-called positive symptoms of schizophrenia (auditory hallucinations, loose associations, inability to concentrate) are due to excess striatal dopaminergic activity, possibly due to elevated numbers of dopaminergic receptors (Seeman et al., 1993). Thus, changes in the intricate balance of dopamine levels in basal ganglia and especially striatum result in problems in executive functions that are either manifested as a reduced ability to initiate actions and to shift sets, when there is too little dopamine, and in an inability to prevent inadvertent shifts in set (*e.g.*, loose associations) when there is excessive dopaminergic activity.

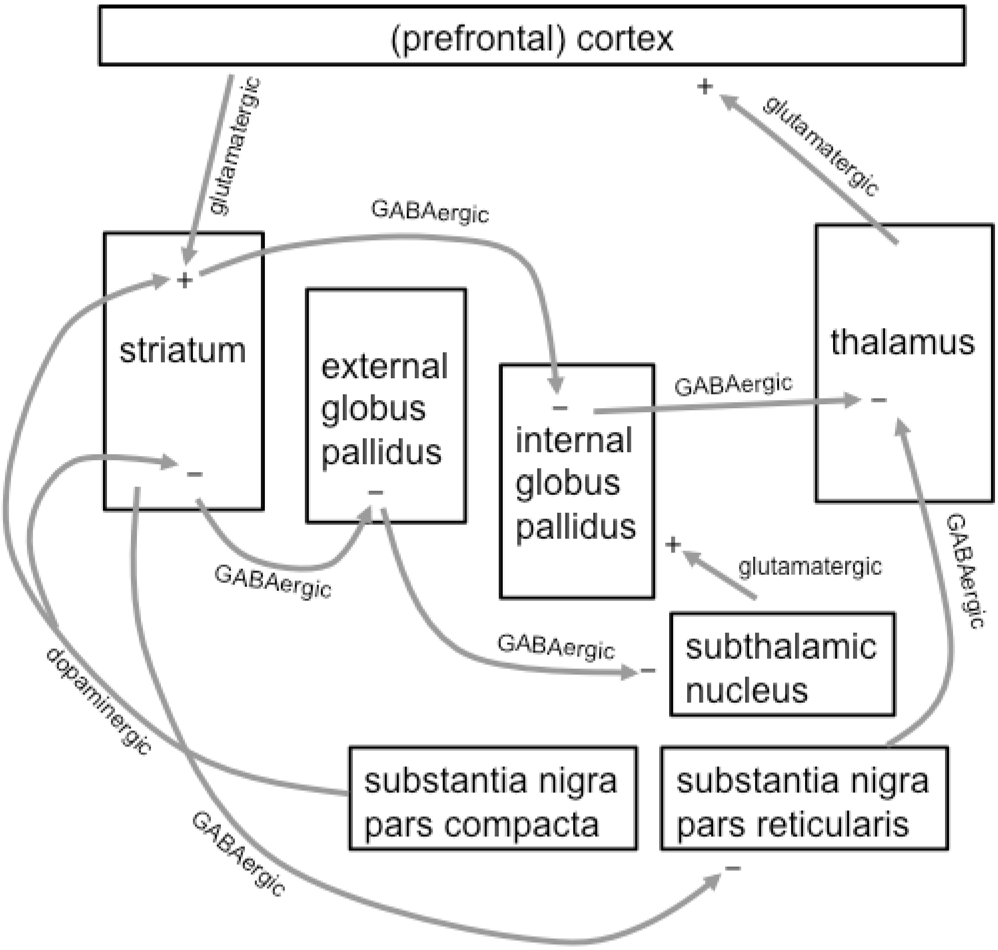


Figure 10-4. Connections of basal ganglia support selection of an action from multiple available alternative actions. Notably, dopaminergic inputs from substantia nigra pars compacta excite striatal neurons that inhibit internal glolus pallidus, which in turn leads to reduction in tonic inhibition exerted by internal globus pallidus neurons onto thalamus. This leads to thalamus sending excitatory inputs to (prefrontal) cortex. The relationship between substantia nigra pars compacta activity and motor activity is not, however, straightforward as electric stimulation of substantia nigra pars compacta in experimental animals does not result in movements, however, in patients suffering from Parkinson's disease, death of dopaminergic nigro-striatal pathways *en masse* result in that the tonic inhibition of thalamic inputs to cortex by internal globus pallidus neurons cannot be lifted and thus initiating movements (as well as shifting of cognitive sets) becomes difficult.

Decision-making

One of the executive functions that prefrontal cortex is also intimately tied with is decision-making. For a recent review studies investigating the neural basis of decision-making, see (Kable and Glimcher, 2009). In non-human primate studies, decision-making has been modulated by varying the probability of receiving food or drink rewards. In humans, decision-making has been most often studied in the context of making economic decisions; experimental subjects have received monetary gains or suffered losses with different probabilities based on their level of risk taking. In such experiments, a number of brain areas have been implicated; valuation has been associated with activity in the ventromedial prefrontal cortex and striatum, and making choices has been associated with activity in lateral prefrontal and parietal areas.

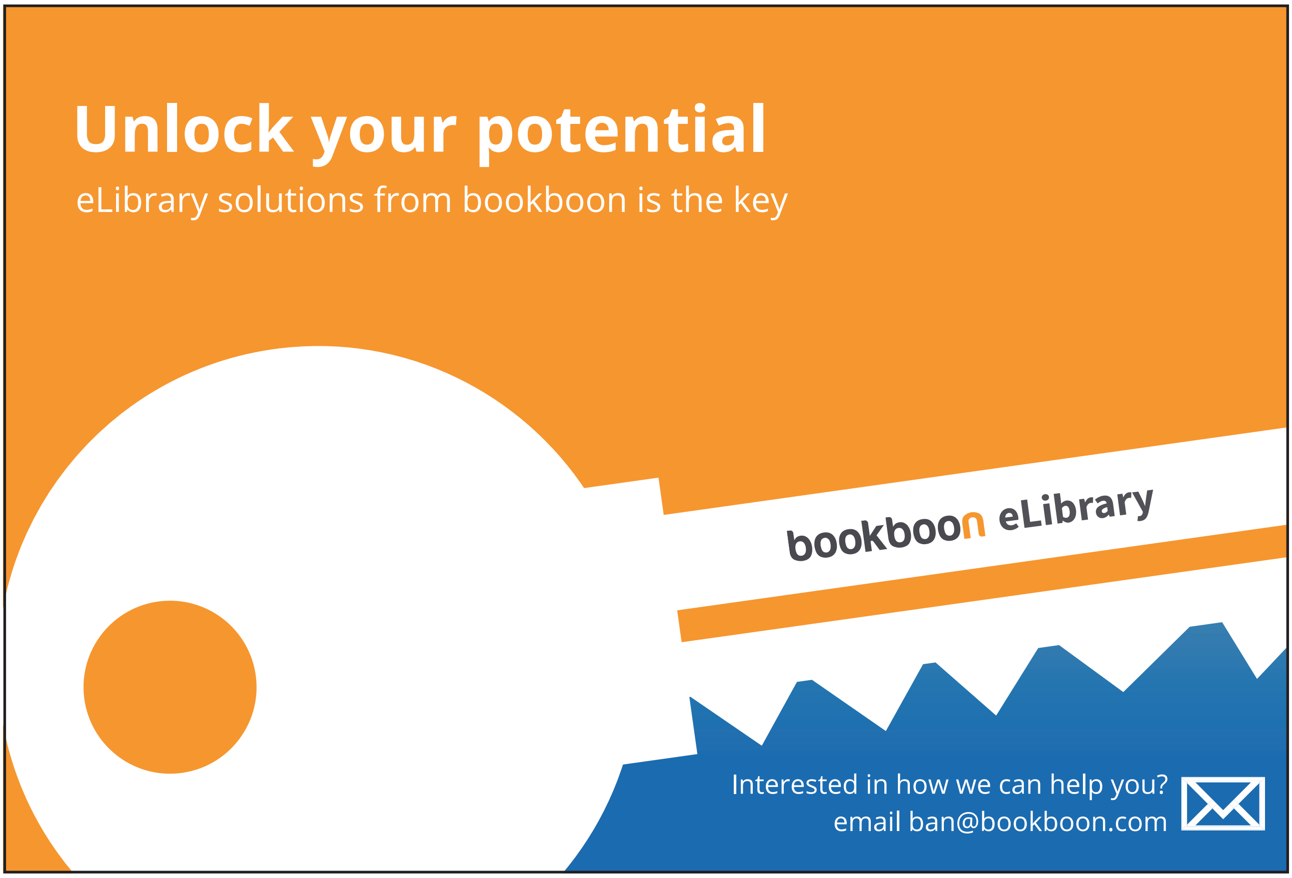
It is good to note that even though monetary rewards/punishments are easiest to implement in a neuroimaging setting, for humans social acceptance and rejection constitute one of the most fundamental of rewards and punishments, respectively.

In addition to economic, social decision-making has been increasingly studied and it seems that the brain mechanisms serving the two are to a large extent shared (for a recent review, see (Rilling and Sanfey, 2011)). Social decision-making is further discussed in Chapter 12 where the neural basis of social cognition will be described.

Role of selective attention and working memory in executive functions

The concepts of selective attention and working memory, introduced in Chapters 6 and 7, respectively, are central also for executive functions, as they involve filtering to, and active manipulation of task/goal-relevant information, in awareness. The role of working memory in executive functions is exemplified by situations where, for instance, one has to mentally plan ahead an important meeting where an important decision needs to be achieved. Relevant information about the personalities and anticipated points of view of the other participants need to be retrieved and actively worked on to form a synthesis of the situation and come up with a relevant negotiation strategy.

Similarly with the other aspects of executive functions, prefrontal cortical areas are also highly important for selective attention and working memory. Dorsolateral prefrontal cortex is an area that has been specifically associated with selective attention and working memory in neuroimaging studies. It is good to keep in mind, however, that both selective attention and working memory depend on more extensive networks of brain areas encompassing, for instance, parietal lobe areas in addition to prefrontal cortical areas (Dove et al., 2000, Lezak et al., 2004).



Schemas and executive functions

Memory representations (schemas) were introduced in Chapter 7 as a mechanism that enables efficient coding of long-term memories; only the most significant details of a given life event need to be encoded as upon retrieval the event is reconstructed based on schema-based general knowledge and the few significant details associated with the event. Similarly, the relevant knowledge, in the form of schemas, is activated when planning ahead and organizing behavior in pursuit of specific goals. Indeed, it is good to bear in mind that the primary purpose of memory is to help make sense out of the world and predict what will happen next due to actions that one is about to take.

An experimental example of how schemas of life events can be imaged in the brains of healthy volunteers is shown in **Figure 10-5** below. During scanning of brain activity with functional magnetic resonance imaging, subjects were processing descriptions of daily life events. Activity elicited when temporally ordering segments of daily life event descriptions (*e.g.*, "get in car", "pour cereal to plate", "wake up", and "leave house") was compared with that elicited by the same texts when subjects were paying attention to the type of font that the text was written with.

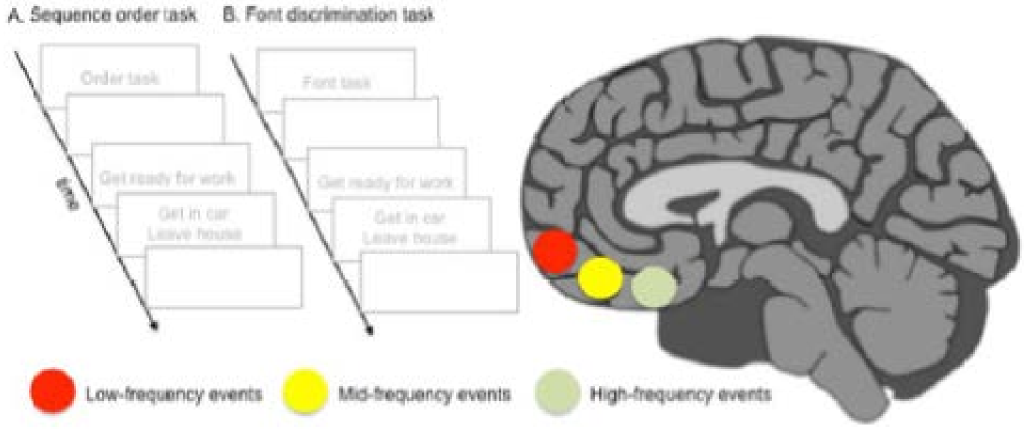


Figure 10-5. Frequency of daily life events shifts the activation focus in an anterior-posterior direction along the ventral-medial prefrontal cortex. Upper left: in the experimental task, the subjects were to put the tasks of a given daily life event into the right order, whereas in a control task, the subjects were to process the fonts that the text was written with. Top right: the degree of familiarity/frequency of a given life event shifted the locus of hemodynamic response in anterior-posterior axis in the ventromedial prefrontal cortex, with less familiar life events activating more anterior areas. This finding provides a good example of a neuroimaging correlate of schemas representing life events. Adapted from (Krueger et al., 2007).

The frequency of occurrence of the life events (*i.e.*, how typical or rare they are) shifted the focus of activation in the anterior-posterior direction along the ventromedial prefrontal cortex, with processing of less frequent/typical life events eliciting responses in more anterior areas (Krueger et al., 2007). In general, there seems to be a posterior-to-anterior hierarchical progression in the prefrontal cortex, with more anterior areas processing events that are more complex and occur over temporally longer timescales. Findings from studies that have inspected how the brain processes information over multiple timescales are described in the following.

Temporal receptive windows and executive functions

Being able to integrate information over time is crucial for perception of life events as they typically unfold over longer timescales. Consider, for instance, a fairly frequent type of event, greeting a person you have not seen in a while. It involves taking the initial eye contact, walking towards each other, saying hellos, shaking hands, and quickly exchanging information about how things are and have been. All this happens over a few tens of seconds or a few minutes, and one has to retain information over the span of the interaction, interpreting facial expressions and what is being said based on the evolving context.

A hierarchy of temporal receptive windows of the brain has been shown in studies that have looked at which brain areas show correlated activity across subjects when they listen to an audio story, either the full forward version, *vs*. when the story is scrambled at paragraph, sentence, or word level. It was observed that the more the temporal structure was preserved, the more anterior frontal lobe areas showed correlated activity across subjects (Lerner et al., 2011). These findings suggested that anterior prefrontal cortical areas are specifically important for integrating information over longer timescales, thus making it possible to interpret individual sentences in the context of the evolving narrative.

Another line of evidence pointing to a hierarchy of temporal receptive windows comes from studies that have looked at specific frequency bands of brain activity in which correlations across subjects are highest in different brain areas when subjects are watching a narrative movie that is composed of information distributed over multiple timescales. Anterior frontal lobe structures seem to exhibit across-subjects correlated activity specifically in lower frequency bands that carry information that evolves more slowly across time (Kauppi et al., 2010). Hierarchy of neuronal adaptation time scales is one potential neural mechanism underlying such findings (*i.e.*, adaptation of neurons for brief periods of time allow the brain to integrate information over time) (Jaaskelainen et al., 2011).

10.3 Neuropsychological tests that measure executive functions

Another way to characterize executive functions is to take a look at behavioral neuropsychological tests that have been developed to measure different aspects of executive functions. It is well known that traditional intelligence quotient tests that are widely used in neuropsychological assessment of cognitive deficits following brain damage often fail to show anything abnormal in patients with prefrontal cortical damage. This lack of sensitivity has resulted in development of specific behavioral tests that do measure executive functions and prefrontal brain damage. In the following, the most widely used tests will be described.

Stroop color-word interference task: inhibiting competing responses

One quite widely used test is a so-called stroop color-word interference test that was introduced in Chapter 2. In the stroop test, which was introduced already in the 1930s, the task of the patient or subject is to suppress a highly automated response and produce a response with less strong association (Stroop, 1935, MacLeod, 1991). Specifically, as shown in **Figure 2-3**, the subject is shown a list of words, names of colors, written with different colors in the Stroop color-word interference task. The task of the subject is then to go as quickly as possible the list through, reciting the colors with which the names of colors are written. As the color with which a given color name is written is different than the color name, the subject needs to suppress the automated response of reading the color names, and instead list aloud the colors with which the text is written. There are a number of variants of the test, but the basic idea of having to suppress competing automated responses is the same in each (MacLeod, 1991).

You can test the Stroop test yourself by creating a list with colored pens and then reading it through, trying to avoid saying out aloud what is written (make the list substantially longer than the example shown in **Figure 2-3**). You should notice a degree of effort when quickly going through the list—this is perfectly normal and indicates that the neural systems responsible for inhibiting task-irrelevant / competing responses are at work. When patients with prefrontal cortical damage try to carry out this task, there are two types of difficulties/deficits. First of all, the performance might be abnormally slow, as compared with the speed that healthy subjects conduct the test. A second, and more indicative, sign of disturbance are rule violations, where the patient occasionally reads out aloud what is written on the list instead of telling what the color is with which the color word is written. These are called interference errors.

Cortical areas involved in different aspects of conflict and competition processing have been identified in recent neuroimaging experiments in healthy subjects (Melcher and Gruber, 2009). Activations related to motor response competition have been reported in the left premotor cortex of right-handed subjects. Semantic incongruence, representing cognitively higher-order conflict, exhibited activation of the anterior cingulate cortex (that is also more generally associated with situations where there is conflicting information that needs to be dealt with), bilateral insula, thalamus, as well as somatosensory cortex. These activations presumably reflect enhanced effort to overcome the interference caused by the strong learned associations. In addition, anterior temporal and rostroventral prefrontal cortical activations have also been described during the Stroop interference task (Melcher and Gruber, 2009). Naturally, these and other similar neuroimaging findings suggest that executive functions depend on more extensive networks of brain areas than just the prefrontal cortical areas, even though the prefrontal cortical areas constitute rather critical nodes in these networks.

Word fluency test

Another test that taps executive function is the so-called word fluency task, where the subjects (or patients) are instructed to produce as many words that they can in short time interval that begin with a designated letter, such as letter s, without producing names of people, and without repeating any given word. Again, this is something that you can easily test yourself. Give yourself two minutes time to come up with as many words as you can beginning with the letter "s". The typical deficits that patients with prefrontal cortical damage exhibit are both lack of words produced (they might manage only few) and rule violations, where they might begin loosely associating the words that they produce with words beginning with another letter (*e.g.*, "short, side, wall, nut, apple…" when the task is to produce as many s-letter words as possible) and/or exhibit perseveration (*e.g.*, "short, side, soon, sun, sun, sun …") (for an in-depth review, see (Lezak et al., 2004)).

In neuroimaging experiments with healthy volunteers, anterior cingulate gyrus, inferior and middle frontal gyri, frontal operculum and cerebellar vermis have been reported to be activated during verbal fluency task performance. There are different types of fluency designs, for instance, the neuroimaging experiment where these activations were observed constrasted categorical fluency (*i.e.*, the task was to produce words from specific categories such as cars, fruits and furniture) with reciting of over-learned sequences (*e.g.*, reciting the months of the year or the letters of the alphabet) (Gurd et al., 2002). Supporting the findings of posterior parietal cortex involvement in voluntary directing of attention (see Chapter 5), the posterior parietal cortex was significantly activated when subjects were rapidly switching between different types of verbal fluency tasks (Gurd et al., 2002).

The Wisconsin card sorting test

The so-called Wisconsin card sorting test (WCST) (short for University of Wisconsin Card Sorting Test) is perhaps the most widely used (and one of the most sensitive) of neuropsychological tests of executive functions (Grant and Berg, 1948, Demakis, 2003). In WCST, the subjects are presented four exemplary cards that differ from each other based on the number, type and color of geometric shapes printed on them (see **Figure 10-6**). The task of the subject is then to guess which of the features, color, shape, or number, is the basis of classification. The experimenter has chosen, without telling the subject, one of these qualities as the matching criteria. Then, the subject begins guessing by matching cards one by one (from a stack of cards) with the four cards on the table in front of him/her.

After each guess, the experimenter tells him/her whether there was a match or whether the subject miss-classified the card. For instance, if the experimenter has decided that color is the matching criteria, then in the example shown in **Figure 10-6**, the card with the single green triangle (which could also be matched with two other cards on the table based on shape or number of items) is a match only when placed in front of the exemplary card with green color (Grant and Berg, 1948). With a few trials and errors the subjects typically learn to categorize/match the cards quite quickly.

It is then that the critical turn occurs in the test; without any advance warning, the experimenter changes the matching criteria to another one, for instance, shape. Then, as the subject who has learned to categorize based on color continues to match with the color criterion, he/she is given negative feedback. When this happens, healthy subjects are typically at first perplexed, but are quickly able to figure out that the matching category has changed, and learn the new categorization principle by trial and error. Then, after a few correct trials the experimenter again changes the matching criteria and the process of re-learning is repeated.

Patients with prefrontal cortical (especially dorsolateral) damage do poorly in WCST (Demakis, 2003). While prefrontal cortex damaged patients attain categories (*i.e.*, learn the categorization rule) more slowly than healthy controls, they are often able to attain the first category. However, they typically experience difficulties when the matching criteria are suddenly changed. After the categorization rule is changed, prefrontal cortical damaged patients commit perseveration errors. Perseveration manifests itself in WCST as repeated attempts to continue matching with the old criteria (in the exemplary case color) even when the patient is getting negative feedback on every trial.

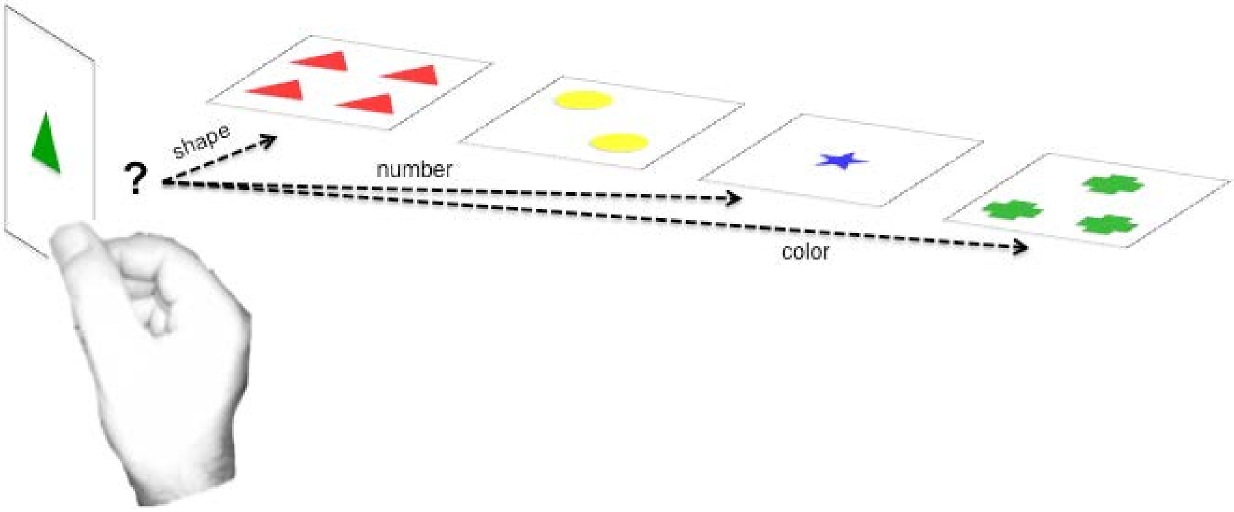


Figure 10-6. The Wisconsin card sorting test. Four cards are first placed in front of the subject, with the cards differing by type, color and number of geometric figures printed on them. Then, the subject draws new cards from a stack and one by one tries to guess what the categorization principle is. The exemplary card shown here in the hand of the subject has one green triangle and thus could be matched with three of the cards on table, based on shape (triangle), number (one figure), and color (green). He/she receives positive or negative feedback based on whether the guess of categorization principle decided by the experimenter is correct. Once the subject gets it right ("attains a category"), the experimenter changes the categorization principle, and the subjects gets negative feedback for matching based on the until-then correct category.

Neuroimaging experiments have delineated patterns of brain activation during performance of the various stages of WCST in healthy volunteers (Monchi et al., 2001). Prefrontal cortical, basal ganglia, and thalamic activations are observed when negative feedback is presented signaling for a need to shift to a new response set. This is quite consistent with the role of prefrontal cortical-basal ganglia-thalamus loop in not only initiation of movements but also shifting of mental sets, as outlined in **Figure 10-4** above. Additionally, there are working-memory related activations that involve dorsolateral prefrontal cortical areas when the subjects are trying to relate feedback, either positive or negative, to preceding events (Monchi et al., 2001).

Observation of behavior in naturalistic settings: multiple errands test

Even specific executive function tests such as the WCST sometimes fail to detect deficits in executive functions following brain damage. Patients have been described who exhibit above normal performance on quantitative neuropsychological tests but are grossly impaired in their ability to organize their lives (Eslinger and Damasio, 1985). Behavioral observation and interviews with the relatives/caretakers of the patients are therefore very important source of additional information, for instance, on whether the behavior of the patient has become more impulsive in everyday life, or whether the patient is lacking initiative. What relatives often describe is that the personality of the patient has somehow changed after the accident. More detailed questioning by the clinician then reveals what type of specific problems the patients manifest in their daily lives.

There are also specific real-life tests that are highly sensitive to executive function deficits. In the so-called multiple errands test, patients are required to carry out multiple fairly simple but open-ended tasks over a 15-30 min time period (Shallice and Burgess, 1991). What is typically found is that patients with deficiencies in executive functions behave in a disorganized manner and end up spending too much time in the test. It has been argued that the problems exhibited by frontal lobe patients on the multiple errands test are due to "an inability to reactivate after a delay previously-generated intentions when they are not directly signaled by the stimulus situation" (Shallice and Burgess, 1991).

Different versions of the multiple errands test and ways to quantify performance errors have been developed, confirming the sensitivity and ecological validity of the test for the measurement of executive function deficits (Dawson et al., 2009). Recently, neuroimaging experiments with highly naturalistic stimulus and tasks environments such as watching movies or playing computer games have been observed to elicit reliable responses in prefrontal cortical areas related to emotions, social cognition, and executive functions (Spiers and Maguire, 2007, Jaaskelainen et al., 2008, Hasson et al., 2010). It is foreseeable that advances in neuroimaging technology and data analysis algorithms will make it possible to develop ecologically valid tests of executive functions also for neuroimaging settings.

10.4 Neurochemistry of executive functions

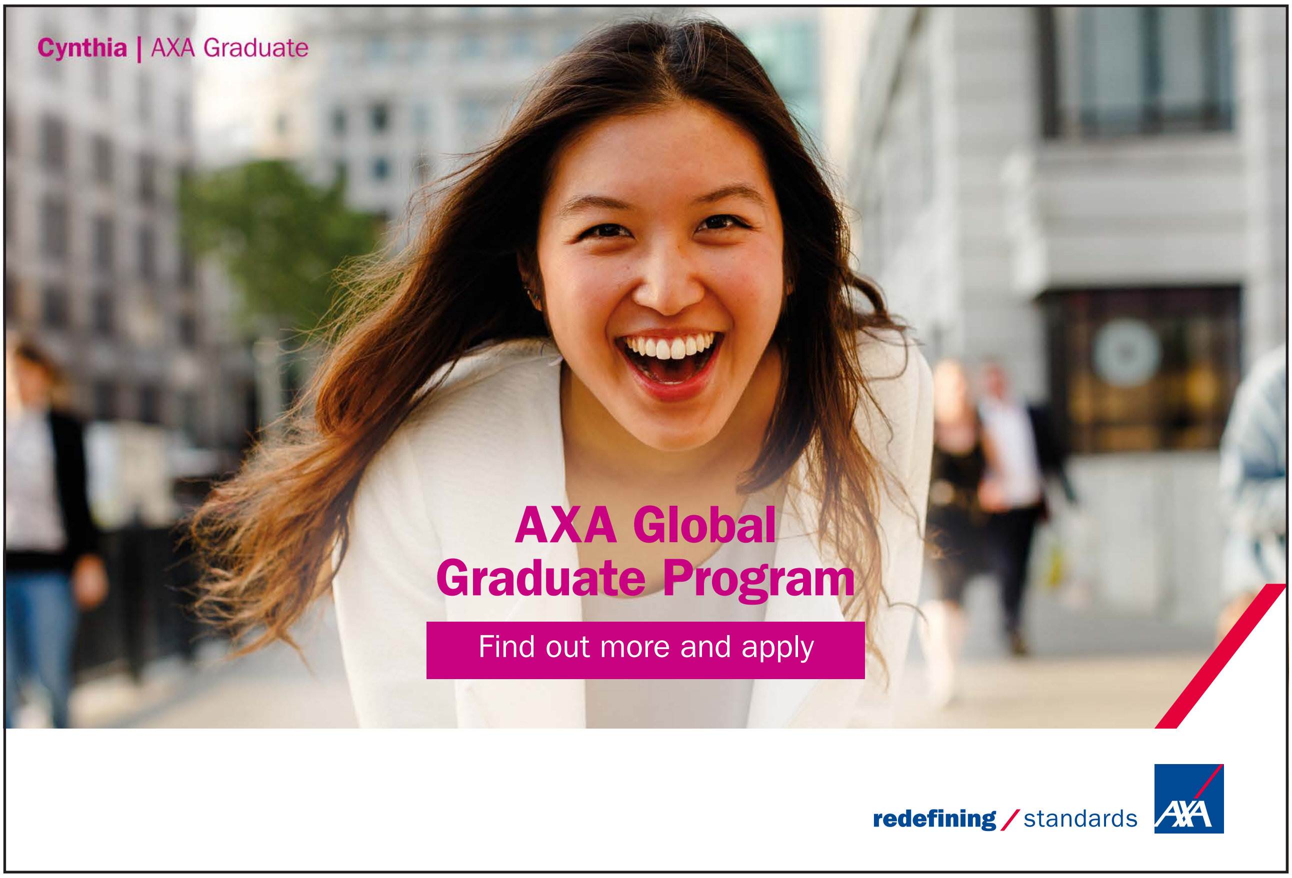
Given the multifaceted nature of executive functions it is not a surprise that there is not a single neurotransmitter that underlies executive functions, but rather an intricate interplay of several neurotransmitter systems enables cognitive functions such as goal-directed behavior and decision making. Given its key role in the loop between prefrontal-cortex, basal ganglia, and thalamus (see **Figure 10-4** above), dopamine is one of the most investigated neurotransmitters in the context of research on executive functions. Indeed, dopamine seems to have a very specific role in choosing between alternative behaviors and cognitive sets.

Role of dopamine in executive functions

Evidence for the involvement of dopamine in executive functions comes from multiple sources. In clinical medicine, patients suffering from Parkinson's disease (due to a massive loss of nigro-striatal dopaminergic neurons) exhibit deficits in executive functions. Specifically, they have problems in ability to shift cognitive set and exhibit perseveration in frontal lobe tests such as WCST and verbal fluency without showing any reduction in intelligence quotient or on tests of recognition memory (Lees and Smith, 1983). It has also been documented that executive function deficits, as well as efficiency of dopamine therapy, depend on the stage of progression of the disease. Early on in the course of the disease the dopaminergic depletion is restricted to striatum but later on spreads to involve the mesocortical dopamine pathway (the dopamine pathways are illustrated in **Figure 6-10**). This leads to reduction in prefrontal cortical dopamine levels, with more progressed cases exhibiting wider array and more severe deficits in executive functions (Leh et al., 2010).

In contrast to Parkinson's disease, patients suffering from schizophrenia have excess dopaminergic activity that is treated with drugs that block/reduce dopamine function. A relatively high proportion of schizophrenia patients suffer from a wide range of cognitive deficits, most notably those involving attention and executive "frontal lobe" (Sullivan et al., 1993) functions (although there are also preserved functions (Gold et al., 2009)) that also correlate with lack of insight that a number of these patients exhibit (Shad et al., 2006). Since there are a number of other brain abnormalities that have been documented in schizophrenia, however, it is not straightforward to associate deficits in executive functions with the dopamine dysfunction in schizophrenia.

Administration of drugs specifically affecting dopaminergic neurotransmission also produces changes in executive functions as assessed with measures of behavioral performance and neural responses recorded with neuroimaging methods. An example of this is provided by a study where healthy volunteers were administered small doses of a dopamine D2 receptor antagonist haloperidol. At the dose used (2 mg) the D2 auto-receptors were blocked and that leads to elevated levels of dopamine in the synaptic cleft. This resulted in reduced distractibility by task-irrelevant sounds and, when distracted, slower recovery times to get back to the primary task. These behavioral findings were accompanied by slower electromagnetic re-orienting responses recorded from the brains of the subjects with magnetoencephalography (Kähkönen et al., 2002).



Other neurotransmitters and executive functions

Given that the neurotransmitters of the brain are tightly intertwined, it can be said that a gross failure in any of the major neurotransmitters is detrimental to goal-directed behavior. There is evidence indicating, for instance, that blocking glutamate function *via* N-methyl-D-aspartate (NMDA) receptors with ketamine adversely affects WCST performance (Krystal et al., 2000). Benzodiazepines, enhancing the effects of gamma-aminobutyric acid (GABA) on $GABA_A$ receptors, and thus inducing hyperpolarization in neurons *via* increased Cl- influx, produce cognitive deficits by sedating subjects. On the other hand, in patients suffering from anxiety the effects of benzodiazepines can be beneficial by increasing the ability of these patients to concentrate on goals/tasks at hand. The intricate interactions between various neurotransmitter systems are exemplified in the connectivity pattern between the prefrontal cortex, nuclei of the basal ganglia, and thalamus shown in **Figure 10-4**. By looking at this figure it is easy to see that the system would be hampered by taking any out any of the neurotransmitters involved.

10.5 Concluding remarks

Executive functions allow one to adopt/generate meaningful goals that span from important longer-term life goals, such as striving to become a medical doctor, through intermediate such as finding out where to apply for medical studies, to very short-term goals, such as getting something to eat when one is hungry. This involves the ability to select a given behavior or cognitive set and suppress other competing ones, switch appropriately between sets (or tasks), and also the ability to take feedback into account when obstacles arise. Failure of executive functions is a fundamental problem in many neurological and psychiatric disorders. The brain area most often implicated in executive functions is the prefrontal cortex.

It is important to keep in mind, however, that the prefrontal cortex is composed of multiple areas that show functional specialization. Furthermore, the prefrontal cortex should be more seen as a central hub in a network of brain areas that govern different aspects of executive functions, rather than an isolated brain structure that underlies executive functions. One example of this is voluntary shifting of attentional focus from one task to another that involves an intricate network of dorsolateral prefrontal cortex and posterior parietal areas. Another good example is the brain network consisting of prefrontal cortical areas, nuclei of basal ganglia, and thalamus, which allows one to adopt new cognitive sets and initiate appropriate movements while suppressing competing ones.

Of the different neurotransmitters, dopamine is the one most intimately and specifically implicated in executive functions, especially in how one is able to choose and shift cognitive sets flexibly. Dopamine should, however, be seen as a central neurotransmitter that influences other neurotransmitters, rather an isolated system that would alone underlie executive functions. Indeed, taking out any of the major neurotransmitters would have disturbing effects on executive functions given that the neurotransmitter systems profoundly interact with and influence one another.

Finally, it is good to bear in mind that executive functions are a rich and multifaceted set of higher-order cognitive abilities. In fact, it is sometimes difficult to draw clear boundaries between executive functions and other higher-order cognitive functions, such as selective attention, working memory, emotions, motivation, and social cognition, which are functions that closely relate to executive functions. Consider, for instance, the role that emotions play in motivating goal-directed behavior. Research findings on emotions and the underlying neural mechanisms will be described in the following chapter.

10.6 References

ARON, A. R., ROBBINS, T. W. & POLDRACK, R. A. 2004. Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences,* 8**,** 170-177.

DAMASIO, H., GRABOWSKI, T., FRANK, R., GALABURDA, A. M. & DAMASIO, A. R. 1994. The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science,* 264**,** 1102-1105.

DAWSON, D. R., ANDERSON, N. D., BURGESS, P. W., COOPER, E., KRPAN, K. M. & STUSS, D. T. 2009. Further development of the multiple errands test: standardized scoring, reliability, and ecological validity of the Baycrest version. *Arhives of Physical Medicine and Rehabilitation,* 90**,** S41-S51.

DEMAKIS, G. J. 2003. A meta-analytic review of the sensitivity of the Wisconsin Card Sorting Test to frontal and lateralized frontal brain damage. *Neuropsychology,* 17**,** 255-264.

DOVE, A., POLLMANN, S., SCHUBERT, T., WIGGINS, C. J. & VON CRAMON, D. Y. 2000. Prefrontal cortex activation in task switching: an event-related potential fMRI study. *Cognitive Brain Research,* 9**,** 103-109.

ESLINGER, P. J. & DAMASIO, A. R. 1985. Severe disturbance of higher cognition after bilateral frontal lobe ablation: Patient EVR. *Neurology,* 35**,** 1731-1741.

GOLD, J. M., HAHN, B., STRAUSS, G. P. & WALTZ, J. A. 2009. Turning it upside down: areas of preserved cognitive function in schizophrenia. *Neuropsychol Rev,* 19**,** 294-311.

GOTHAM, A. M., BROWN, R. G. & MARSDEN, C. D. 1988. 'Frontal' cognitive function in patients with Parkinson's disease 'on' and 'off' levodopa. *Brain,* 111**,** 299-321.

GRANT, D. A. & BERG, E. A. 1948. A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a weigl-type card-sorting problem. *Journal of Experimental Psychology,* 38**,** 404-411.

GURD, J. M., AMUNTS, K., WEISS, P. H., ZAFIRIS, O., ZILLES, K., MARSHALL, J. C. & FINK, G. R. 2002. Posterior parietal cortex is implicated in continous switching between verbal fluency tasks: an fMRI study with clinical implications. *Brain,* 125**,** 1024-1038.

HARLOW, J. M. 1848. Passage of an iron rod through the head *Boston Med Surg J,* 39**,** 389-393.

HASSON, U., MALACH, R. & HEEGER, D. J. 2010. Reliability of cortical activity during natural stimulation. *Trends Cogn Sci,* 14**,** 40-48.

JAASKELAINEN, I. P., AHVENINEN, J., ANDERMANN, M. L., BELLIVEAU, J. W., RAIJ, T. & SAMS, M. 2011. Short-term plasticity as a neural mechanism supporting memory and attentional functions. *Brain Research,* 1422**,** 66-81.

JAASKELAINEN, I. P., KOSKENTALO, K., BALK, M. H., AUTTI, T., KAURAMÄKI, J., POMREN, C. & SAMS, M. 2008. Inter-subject synchronization of prefrontal cortex hemodynamic activity during natural viewing. *The Open Neuroimaging Journal,* 2, 14-19.

JEANNEROD, M. 1995. Mental imagery in the motor context. *Neuropsychologia,* 33, 1419-1432.

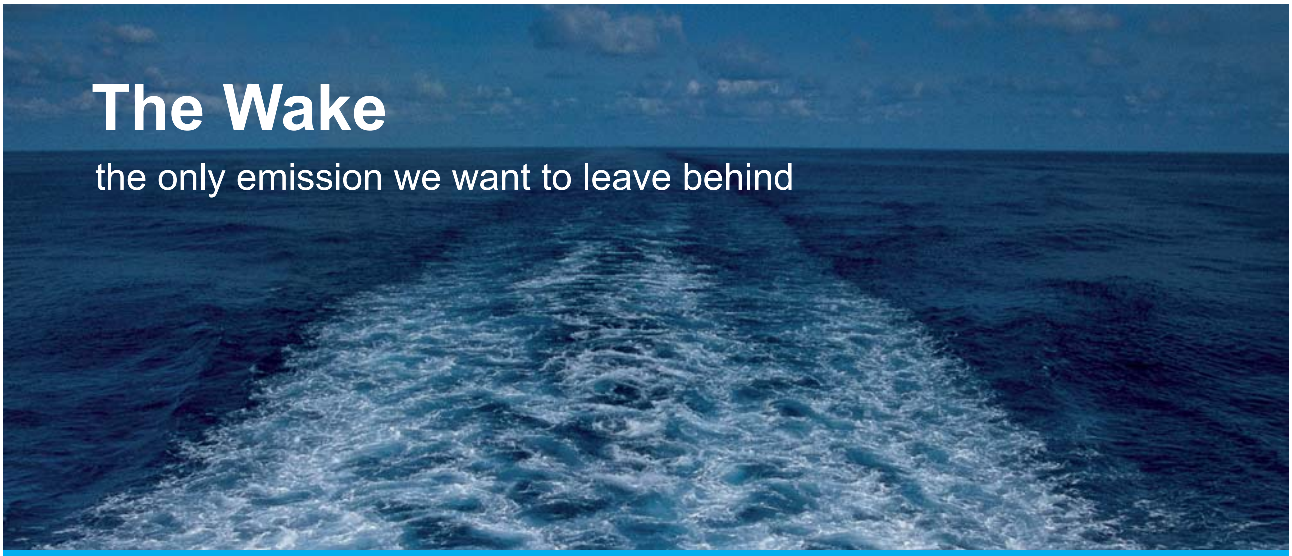
KABLE, J. W. & GLIMCHER, P. W. 2009. The neurobiology of decision: consensus and controversy. *Neuron,* 63, 733-745.

KÄHKÖNEN, S., AHVENINEN, J., PEKKONEN, E., KAAKKOLA, S., HUTTUNEN, J., ILMONIEMI, R. J. & JÄÄSKELÄINEN, I. P. 2002. Dopamine modulates involuntary attention shifting and reorienting: an electromagnetic study. *Clinical Neurophysiology,* 113, 1894-1902.

KAUPPI, J. P., JAASKELAINEN, I. P., SAMS, M. & TOHKA, J. 2010. Inter-subject correlation of brain hemodynamic responses during watching a movie: localization in space and frequency. *Front Neuroinformatics,* 4, 5.

KRUEGER, F., MOLL, J., ZAHN, R., HEINECKE, A. & GRAFMAN, J. 2007. Event frequency modulates the processing of daily life activities in human medial prefrontal cortex. *Cereb Cortex,* 17, 2346-2353.

KRYSTAL, J. H., BENNETT, A., ABI-SAAB, D., BELGER, A., KARPER, L. P., D'SOUZA, D. C., LIPSCHITZ, D., ABI-DARGHAM, A. & CHARMEY, D. S. 2000. Dissocation of ketamine effects on rule acquisition and rule implementation: possible relevance to NMDA receptor contributions to executive cognitive functions. *Biol Psychiatry,* 47, 137-143.



LEES, A. J. & SMITH, E. 1983. Cognitive deficits in the early stages of Parkinson's disease. *Brain,* 106**,** 257-270.

LEH, S. E., PETRIDES, M. & STRAFELLA, A. P. 2010. The neural circuitry of executive functions in healthy subjects and Parkinson's disease. *Neuropsychopharmacology,* 35**,** 70-85.

LERNER, Y., HONEY, C. J., SILBERT, L. J. & HASSON, U. 2011. Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J Neurosci,* 31**,** 2906-2915.

LEZAK, M. D., HOWIESON, D. & LORING, D. W. 2004. *Neuropsychological assessment (4th ed.),* Oxford, Oxford University Press.

LOMBARDI, W. J., ANDREASON, P. J., SIROCCO, K. Y., RIO, D. E., GROSS, R. E., UMHAU, J. C. & HOMMER, D. W. 1999. Wisconsin Card Sorting Test performance following head injury: dorsolateral fronto-striatal circuit activity predicts perseveration. *Journal of Clinical and Experimental Neuropsychology,* 21**,** 2-16.

MACLEOD, C. M. 1991. Half a century of research on the stroop effect: an integrative review. *Psychological Bulletin,* 109**,** 163-203.

MELCHER, T. & GRUBER, O. 2009. Decomposing interference during Stroop performance into different conflict factors: an event-related fMRI study. *Cortex,* 45**,** 189-200.

MONCHI, O., PETRIDES, M., DOYON, J., POSTUMA, R. B., WORSLEY, K. & DAGHER, A. 2004. Neural bases of set-shifting deficits in Parkinson's disease. *Journal of Neuroscience,* 24**,** 702-710.

MONCHI, O., PETRIDES, M., PETRE, V., WORSLEY, K. & DAGHER, A. 2001. Wisconsin card sorting revisited: distinct neural circuits participating in different stages of the task identifed by event-related functional magnetic resonance imaging. *Journal of Neuroscience,* 21**,** 7733-7741.

ÖNGUR, D., FERRY, A. T. & PRICE, J. L. 2003. Architectonic subdivisions of the human orbital and medial prefrontal cortex. *Journal of Comparative Neurology,* 460**,** 425-449.

RILLING, J. K. & SANFEY, A. G. 2011. The neuroscience of social decision-making. *Annual Review of Psychology,* 62**,** 23-48.

SCHACHAR, R., TANNOCK, R., MARRIOTT, M. & LOGAN, G. 1995. Deficient inhibitory conrtol in attention deficit hyperactivity disorder. *J Abnorm Child Psychol,* 23**,** 411-437.

SEEMAN, P., GUAN, H.-C. & VAN TOL, H. H. M. 1993. Dopamine D4 receptors elevated in schizophrenia. *Nature,* 365**,** 441-445.

SHAD, M. U., TAMMINGA, C. A., CULLUM, M., HAAS, G. L. & KESHAVAN, M. S. 2006. Insight and frontal cortical function in schizophrenia: a review. *Schizophrenia Res,* 86**,** 54-70.

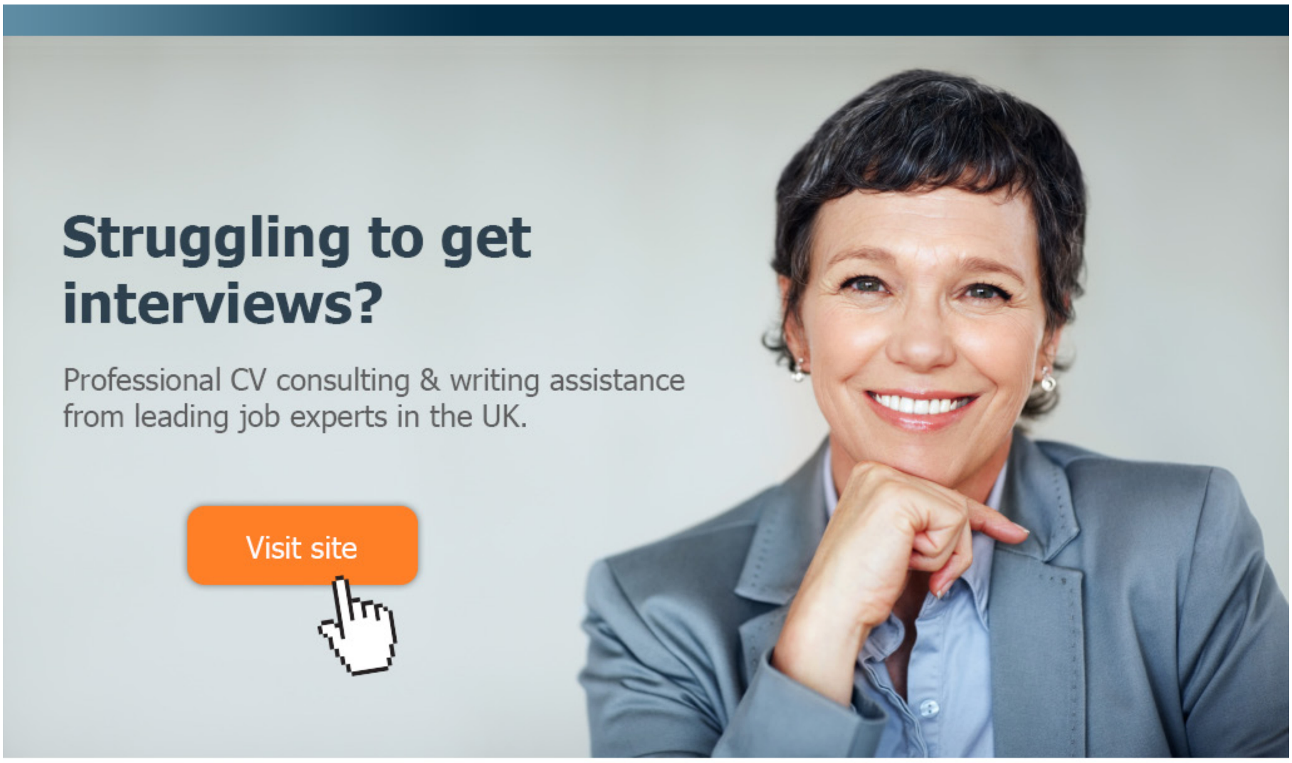
SHALLICE, T. & BURGESS, P. W. 1991. Deficits in strategy application following frontal lobe damage in man. *Brain,* 114, 727-741.

SHUE, K. L. & DOUGLAS, V. I. 1992. Attention deficit hyperactivity disorder and the frontal lobe syndrome. *Brain and Cognition,* 20, 104-124.

SPIERS, H. J. & MAGUIRE, E. 2007. Decoding human brain activity during real-world experiences. *Trends Cogn Sci,* 11, 356-365.

STROOP, J. R. 1935. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology,* 18, 643-662.

SULLIVAN, E. V., MATHALON, D. H., ZIPURSKY, R. B., KERSTEEN-TUCKER, Z., KNIGHT, R. T. & PFEFFERBAUM, A. 1993. Factors of teh Wisconsin card sorting test as measures of frontal-lobe function in schizophrenia and in chronic alcoholism. *Psychiatry Res,* 46, 175-199.



11 Emotions

Imagine a situation where a young adult has to approach an attractive person of the opposite sex that he (or increasingly nowadays she) wants to ask out for a date. On one hand intense crush propels him/her forward and on the other hand there is severe nervousness due to fears of getting turned down. For a young person, getting to go out with the person one has a crush on can be gravely important and there are intense autonomic nervous system / bodily reactions such as heart beating faster, voice cracking, and palms sweating. Imagining a situation such as this, one can appreciate the significance of emotions in bringing about motivation to select and pursue goals in life. Indeed, one is constantly being pushed towards certain things and repelled away from certain others due to anticipation/prediction of rewarding and unwanted outcomes, respectively.

Despite the significance of emotions for human cognition, it was still not long ago that cognition was surprisingly widely thought to be separable from emotions. It was also widely held that emotions would be too difficult and elusive a phenomenon to study rigorously in laboratory or neuroimaging settings. As a result of this, the bulk of cognitive neuroscience research has in the past focused on studying cognitive and perceptual processes. Research on emotions and the underlying neural mechanisms has only begun to flourish relatively recently. This surge in interest towards emotions is highly important, given that emotions are the core driving force that builds motivation and influences goal directed behavior; in many ways cognition and perception are never quite separable from emotions. To begin the description of cognitive neuroscience of emotions, different ways that emotions have been defined will be introduced first.

11.1 Definitions of emotions

When observing the behavior of lower animals, it becomes obvious that emotions are behaviorally manifested in two types of behavior: the animals are attracted towards certain objects in their environment such as food, sweets, or water, and withdraw or flee away from others, such as threatening predators. In humans, of course, emotions are manifested in a more complex manner than (at least what can be observed) in animals, and there are countless words for different flavors of emotions and feelings, such as angry, longing, sad, depressed, melancholic, happy, overjoyed, in love etc. This has motivated researchers to attempt to define "core" or "basic" emotional states.

Certain well established definitions of emotions go even further in the attempt to reduce dimensionality of emotion definitions, and are not all that different from what is seen in emotional behaviors of animals. In these models, where emotions are placed on two dimensions called valence and arousal, emotions are seen as helping humans to adapt to the environment as optimally as possible, by allowing determination of whether something is good or bad, and how significant it is. To complement these views, there are separate definitions for aesthetic emotions as the definitions of basic/dimensional emotions have failed to capture the breadth of feelings that occur when, for instance, listening to music or watching movies (Zentner et al., 2008). The models of basic emotions and dimensional emotion models are outlined in more detail in the following.

Basic emotions: anger, disgust, fear, happiness, sadness and surprise

In search for human core emotions, researchers have shown sets of photographs displaying faces expressing various emotions to subjects from different cultures. This way, it has been possible to show that there are at least six emotional expressions that are relatively consistent across cultures (Ekman et al., 1969). These emotions are anger, disgust, fear, happiness, sadness, and surprise (see **Figure 11-1**). The fact that these "basic" emotional expressions are present across cultures has been taken to suggest that they are based on human biology rather than shaped by culture. The number of basic emotions has been a topic of some debate. There is evidence suggesting that contempt would be a distinct, culturally universal, seventh basic emotion (Ekman and Heider, 1988). On the other hand, it has been sometimes argued that surprise would not be a genuine basic emotion, as it is easily confused with fear. Nonetheless, the theory of basic emotions has been a very influential one in cognitive neuroscience.

Figure 11-1. The six culturally invariant basic emotions: happiness, surprise, fear, anger, disgust, and sadness. The fact that these emotional expressions can be found across different cultures suggests that they are biologically based rather than the result of cultural influences on development. Adapted from (Ekman, 1999).

A recent meta-analysis compared across 105 functional magnetic resonance imaging studies where in the brain hemodynamic responses are elicited to pictures of faces expressing the six basic emotions (Fusar-Poli et al., 2009). The results of this meta-analysis are displayed in **Figure 11-2** below. It seems that there are both areas of shared and specific activations. Happy, fearful, and sad faces activated the amygdala, with the most robust responses being elicited by fearful faces. Insula, on the other hand, exhibited specific sensitivity to expressions of disgust and anger, with strongest responses to disgusted faces. In addition to these more specific responses, there were responses to emotional faces in visual, temporo-parietal, and prefrontal cortical areas, but also in putamen and cerebellum.

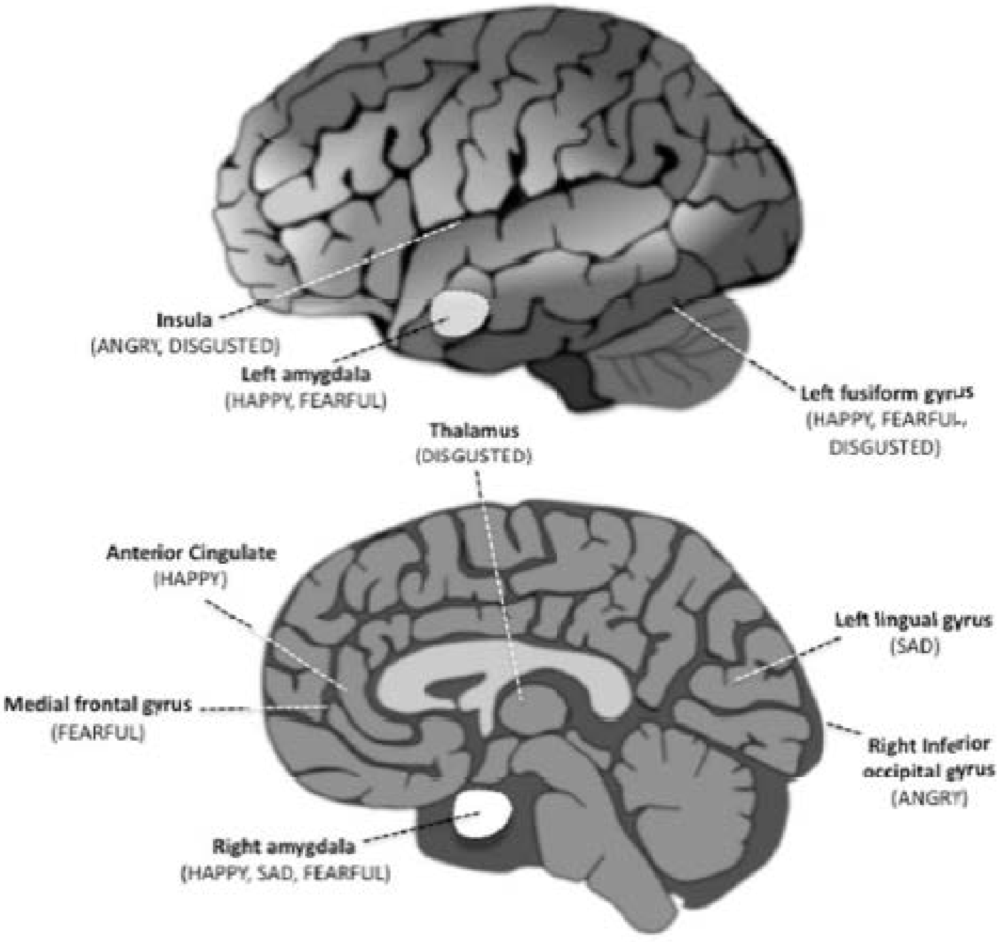


Figure 11-2. Loci of brain hemodynamic responses to six basic emotions. See text for details. After (Fusar-Poli et al., 2009).

A meta-analysis of this type, comparing results across more than a hundred individual functional neuroimaging experiments, is highly useful given that the limitations in signal-to-noise in any given individual experiment might result in false positive findings (*i.e.*, finding brain responses that are not real ones, but rather a chance findings due to experimental noise). It has to kept in mind, however, that the problem of false negative findings (*i.e.*, failing to see brain responses that are really there) is not removed by meta-analyses if, for instance, the vast majority of studies included in the meta analysis have used stimuli that fail to elicit emotions robustly enough in experimental subjects, or have used neuroimaging methods that are less sensitive than those that are presently available (due to the very rapid progression in non-invasive neuroimaging method development that constantly increases the sensitivity of these methods to weak signals related to emotional responses).

Dimensional theories of emotions: valence and arousal

In addition to the model of basic emotions, there are so-called dimensional models of emotion (Schlosberg, 1954, Russell, 1980, Russell and Barrett, 1999). In these models, emotions are modeled using two continuums called valence and arousal. Valence refers to evaluation of emotion-eliciting stimuli/events on a bipolar dimension ranging from repelling/bad to attractive/good. Arousal on the other hand, relates to the intensity of experienced emotions and reflects the degree of autonomic nervous system responses in a given emotional situation. Emotional valence and arousal dimensions can also be used to describe the basic emotions; happiness is a state of positive valence, sad can be characterized as negative valence with low arousal, anger and fear is characterized as negative valence with high arousal, surprise as relatively neutral valence with high arousal, and disgust as negative valence and low arousal.

Evidence for more than two emotional dimensions has been presented (Schlosberg, 1954). There is also behavioral evidence suggesting that instead of a single bipolar dimension running from negative to positive valence could be better described by separate positive and negative valences (Cacioppo and Berntson, 1994). These could, presumably, be also present simultaneously in certain situations such as when watching movies (Jaaskelainen et al., 2008). This state, where one is simultaneously experiencing positive and negative feelings, is referred to as ambivalence.

Recent functional magnetic resonance imaging studies have mapped brain activity correlates of valence and arousal. In one such study, both bipolar (*i.e.*, linear dependency) and separate negative-to-positive (*i.e.*, second-order U-shaped/ inverted U-shaped relationships) valence dimensions were used in modeling hemodynamic responses during watching of emotional pictures (see **Figure 11-3**). It was observed that several brain areas, including dorsomedial and dorsolateral prefrontal cortical areas, anterior cingulate gyrus, and insula, exhibited second-order correlations, thus suggesting that negative and positive valence are processed relatively independently in these brain areas (Viinikainen et al., 2010). Lending support to this viewpoint, non-human primate studies have also demonstrated that there are neurons in the amygdala that separately respond to positive and negative valence stimuli (Paton et al., 2006).

Whilst the categorical (basic emotion) and dimensional models of emotional experiences allow one to effectively classify emotional experiences, these models *per se* do not respond to the question of how the emotional experiences come about. For instance, if one is walking in a dark alley during late hours of the night and suddenly sees two men clad in rugged clothes wielding an axe and a handgun, what is the cascade of psychological and physiological events that result in behavioral responses (running away) and feelings of fear? Research that has provided answers to this intriguing question will be introduced next.

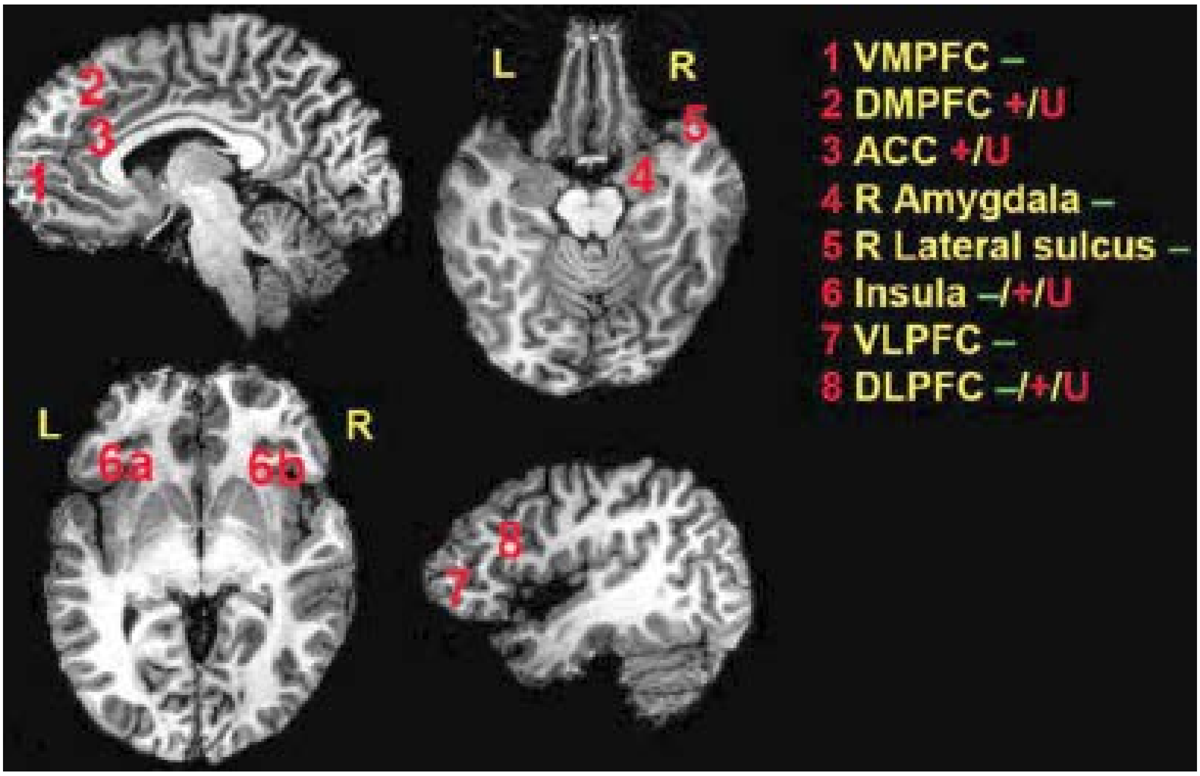


Figure 11-3. Loci of brain regions that correlate with emotional valence of pictorial stimuli. The green and red colored "–" and "+" signs denote linear correlations with negative and positive valence, respectively, whereas the red "U" sign denotes a second-order relationship. (Abbreviations of anatomical area names: VMPFC = ventromedial prefrontal cortex, DMPFC = dorsomedial prefrontal cortex, ACC = anterior cingulate cortex, insula, VLPFC = ventrolateral prefrontal cortex, DLPFC = dorsolateral prefrontal cortex; adapted from (Viinikainen et al., 2010)).

11.2 Emotions result from contextual interpretation of autonomic nervous system responses

The question of how one comes to experience emotions has been one of the most central questions in emotion research. In the early 19th Century, a so-called James-Lange theory of emotions postulated that experiencing emotions are the result of autonomic nervous system reactions (James, 1890). For instance, encountering a lion when on foot in a jungle first elicits the response to flee, followed by autonomic nervous system reactions that generate emotions. The brain then interprets, given also the context, the response such as running away and bodily reactions such as heart beating faster, resulting in feeling of specific emotions.

In the James-Lange theory, darting away from the approaching lion happens first, instead of emotions elicited when seeing the lion causing one to run away. At its time, the theory was seen as highly counterintuitive one. Furthermore, the theory received criticism given that attempts to find emotion-specific autonomic system responses failed; rather it was consistently observed in early 20th Century studies of autonomic nervous system responses that "*the same visceral changes occur in very different emotional states and in non-emotional states*" (Cannon, 1929).

What the James-Lange theory was insightful in is that emotions are elicited with a lag that would have made it very difficult for our ancestors to avoid being eaten by the predators if they would have started running as the result of emotions experienced by them. Thus, it is clear that there is an initial response that causes behavioral reactions, followed by consciously experiencing emotions, at least in case of sudden reactions to threats. But how do the experience of specific emotions (or feelings) then come about given that autonomic system reactions are rather unspecific?

One answer to this question was provided by an ingenious study that Schachter and Signer conducted in the early 1960s. In this study, volunteers were led to believe they had come to a drug study testing the effect of "Suproxin" on vision. In reality, the subjects received injections of either adrenaline or saline. Some of the subjects were then informed of the bodily reactions that they should shortly experience due to the injection and others were told that no side effects were to be anticipated. The subject was then left waiting in a room with (unbeknownst to the subject) a professional stooge, who started behaving either in a jubilant way or expressed increasingly strong anger.

It was hypothesized by the experimenters that the, rather unspecific, autonomic system reactions caused by the adrenalin injection are interpreted by the subjects as distinct feelings given the emotional cues. The results supported this hypothesis; subjects who were informed of the autonomic nervous system effects of the injection attributed their bodily changes to the injection, but those who were uninformed felt happy when they were in the same room with the happy stooge and angry when they were accompanied by the stooge who became increasingly angry (Schachter and Singer, 1962). Based on these results it was concluded that experienced emotional states result from changes in physiological arousal and cognitive interpretation of the arousal given contextual and other cues.

11.3 Induction of emotions in experimental settings

One of the challenges in empirical studies of emotions is how to elicit emotions robustly and reliably enough to allow, for instance, neuroimaging of emotional responses in the brain. This especially concerns studies that attempt to address the neural basis of experiencing emotions (for example it can be argued whether pictures of snakes and spiders cause

robust emotional experiences in experimental subjects), but also to some extent studies that map brain responses during perception/classification of emotional stimuli. Different methods that have been utilized in neuroimaging studies to elicit emotions are described in the following.

Induction of emotional states by mentalizing/imagery

Asking experimental subjects to induce upon themselves specific emotional states during neuroimaging by the means of mental imagery is one of the methods that have been used in eliciting emotions during neuroimaging. Often subjects are helped in mood induction by presenting emotional pictures or narratives (Schneider et al., 1997, Onton and Makeig, 2009). During the process of mood induction, the subjects might imagine/recall certain emotional events. Naturally, the process is not instantaneous but it takes a while for the subjects to attain the desired emotional state.

For some subjects, mood induction by mental imagery is easier than for others, for instance, professional actors use mood induction on a routine basis in their work. Naturally, there might be an array of other responses than those associated with the emotional experiences *per se*, but one way to control for the confounding effects is contrasting/comparing emotional states between each other, and correlating the self-reported robustness of experienced emotions with the observed brain activity effects. Nonetheless, mood induction by mental imagery is a highly effective method for eliciting emotions that allow measurement of the associated brain responses.

Presenting emotional stimuli to volunteers

Presenting emotional stimuli to experimental subjects is perhaps the method that has been most often utilized when studying neural correlates of emotional processing. There are a variety of stimulus types, ranging from still pictures, through narrated audio stories and music to video clips and even entire movies. The international affective picture system (IAPS) is one example of rigorously collected set of emotion-eliciting stimuli (Lang et al., 2005). In the IAPS, there are hundreds of emotional pictures that have been rated for valence and arousal in a large population of healthy subjects.

Pictures of small babies are good examples of positive valence pictures, whereas mutilated bodies are rated as highly negative. Calming pictures of landscape are rated low in arousal, whereas nude pictures tend to be rated as highly arousing by experimental subjects. The brain structures showing significant correlations with emotional valence are shown in **Figure 11-3** above. Pictures are not, however, the most effective way to elicit robust emotions in experimental subjects. In a study comparing different stimulus types, it was observed that narrated stories and movies elicited most robust emotions out of the stimulus types tested in healthy volunteers (Westermann et al., 1996).

Use of rewards and punishments to elicit emotions and motivation

Use of rewards and punishments has been a central means to study emotions in experimental animals. For instance, much of what is known about how the amygdala codes stimuli that predict aversive events, and triggers autonomic nervous system reactions following presentation of them, has been due to work in animals. Specifically, in these studies, experimental animals are presented neutral stimuli, such as a tone, followed by an aversive electric shock. It has been observed that the lateral nucleus of amygdala pairs the neutral stimulus with the aversive stimulus through the mechanism of long-term potentiation (Rogan et al., 1997). Then, when the neutral stimulus is presented again, the lateral nucleus recognizes it and triggers the autonomic nervous system / emotional responses *via* the central nucleus of the amygdala.

This is schematically illustrated in **Figure 11-4** below. In human studies it has been also shown that the amygdala widely modulates the cortical processing of stimuli, as will be described in more detail later in this chapter.

In addition to using punishment, rewarding stimuli have been used to study approach behavior and development of addictions in animal models. In these studies, tasty liquids and foods are typically used as the reinforcing stimulus; the reward is given whenever the animal produces the to-be-reinforced behavior. Dopaminergic neurons, especially in the ventral tegmental area of the brain, have been especially observed to be associated with reinforcement of behavior by rewarding stimuli. Specifically, it has been shown that dopamine activity codes for the expectation value of reward (*i.e.*, "reward prediction error").

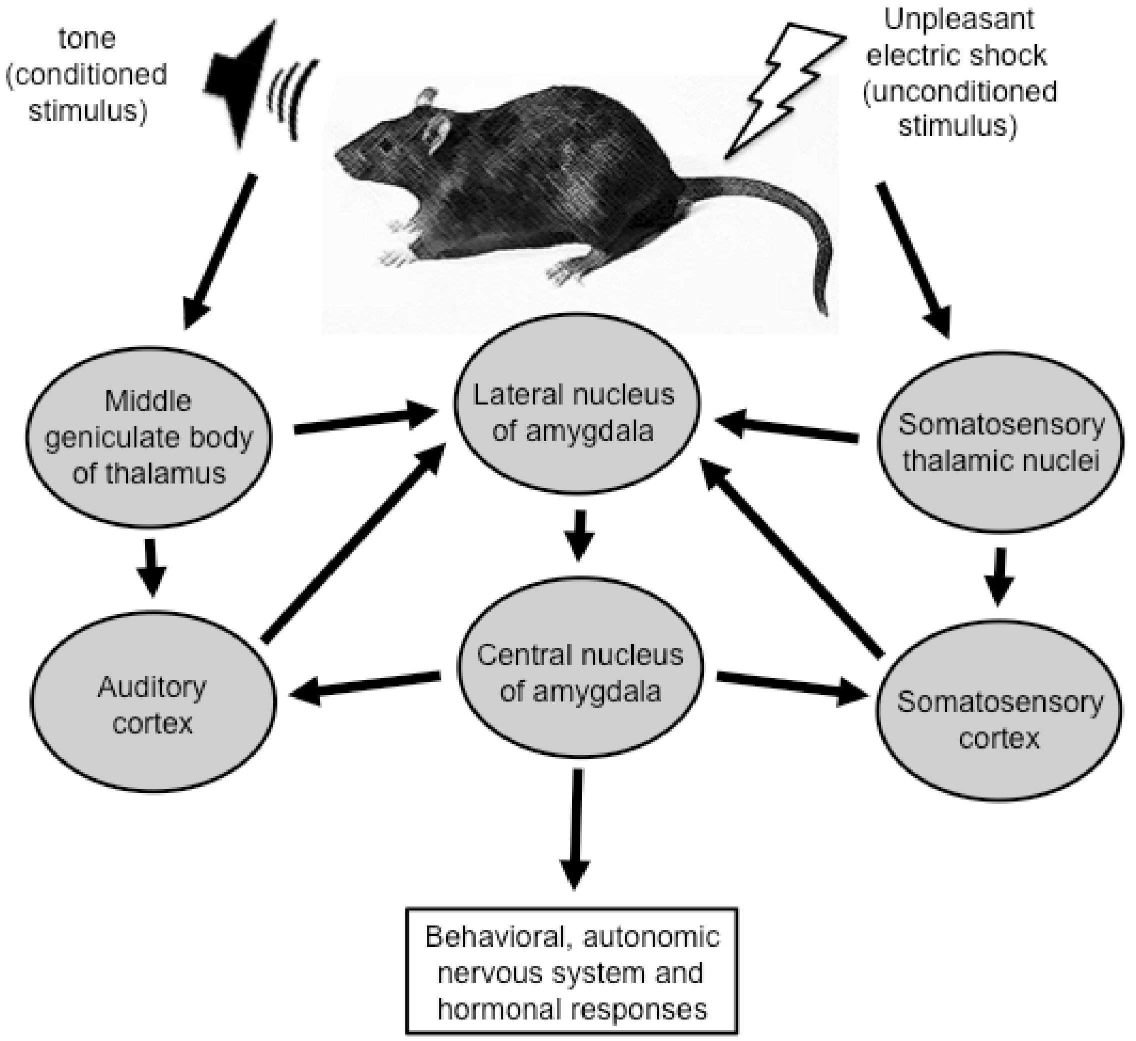


Figure 11-4: Coding of emotional stimuli by amygdala enable quick emotional responses to important stimuli. Here, an experimental setup is shown where a neutral tone is played followed by an unpleasant electric shock. The tone is then coupled with the electric shock in the lateral nucleus of the amygdala. When the tone is subsequently presented, the lateral amygdala neurons that react to the tone after the coupling then trigger, *via* the central nucleus of the thalamus, behavioral, autonomic nervous system and hormonal responses that are central constituents of emotions. It is noteworthy that there are direct thalamic connections to amygdala that also form the basis of this learning. Thus, emotional stimuli access the circuitry very quickly *via* thalamus, without having to be first processed cortically. Furthermore, amygdala does widely influence cortical processing of emotion-eliciting stimuli.

When a reward is for the first time presented with a certain action, dopamine neurons fire after the reward. Then, once the association between the action and the reward is established, the dopamine neurons fire vigorously when the stimulus leading to the reward is presented, but not to the reward itself. Furthermore, if a reward is not given, the dopamine neurons again fire (Schulz, 2000). These observations suggest that dopamine neurons code for the expected reward value of a given stimulus and thus help the animal select what actions to take. Along these lines, it has also been interpreted that dopamine firing underlies "wanting of something", instead of being the neural correlate of feeling pleasure (Berridge and Robinson, 1998), something that the opioidergic system (that is closely coupled with the dopamine system) probably underlies (Pecina and Berridge, 2000).

While animal models have been highly useful in studying certain aspects of emotions and motivation and their underlying neural mechanisms there are, naturally, shortcomings. For one, it is always somewhat questionable to what extent the emotional experiences of the animals match those of humans. Given that one cannot ask the experimental animals how they feel, one has to assume that the stimuli elicit emotions, given the presence of, for instance, approach, withdrawal, or startle behavior. On the other hand, measuring emotions is not trivial in human subjects either and there are a variety of methods that have been used to assess emotions. These will be described in the following.

11.4 Methods for assessing emotions

The methods used in assessing emotions include asking subjects directly how they feel using various self-report questionnaires, using indirect methods to probe emotional responses to stimuli, observing the reactions of subjects including measurement of facial muscle activity, as well as measuring the responses of the autonomic nervous system. Increasingly, it is also possible to use neuroimaging methods to assess emotional reactions of experimental subjects. Naturally, by using more than one method, the accuracy of assessment increases. On the other hand, time constraints in empirical work are often strenuous as subjects' time is valuable. Subjects also get tired if the experimental procedure takes excessive amounts of time, which reduces the quality of neuroimaging data that is obtained. Thus, in practice, most often only a few selected measures are utilized in a given study.

Self-report questionnaires of emotional states

By far the simplest and in many ways the most robust and reliable method of assessing experienced emotions is asking the human subjects directly with different types of questionnaires how they feel (Barrett et al., 2007). In Likert-scale questionnaires, named after the inventor Renses Likert, the subjects are asked to rate their emotions on either discrete steps or continuums ranging from very low to very high degree of experienced emotion. The visual analogue scale is an example of continuous scale used in obtaining self-reports of emotions. In the visual analogue scale, the subject is presented with a line of certain length and is instructed to mark on the line, for instance, how intense the emotion is that they are experiencing, with very low typically on the extreme left and very high on the extreme right.

What the subjects are specifically asked to rate depends of course on the experimental setup, the research questions that are being assessed, and the theoretical framework. In some experiments, the intensity of the basic emotions is what the experimenter is interested in assessing. In others, the subject might be asked to rate his/her valence and arousal during various stages of the experiment or in response to emotional stimuli. In studies that address the neural basis of higher-order aesthetic emotions, such as pleasant melancholy brought about by listening to classical music, the questionnaires are more fine-grained.

Self-reports measuring emotional response trait characteristics

Questionnaires are also often utilized to assess emotional trait characteristics of the subjects, for instance, the degree of empathy of the subjects. Empathy can be dissociated to subclasses of cognitive and emotional empathy (Davis, 1983). The types of questions that have been used to address cognitive empathy include questions that assess perspective-taking ability (*e.g.*, "I sometimes try to understand my friends better by imagining how things looks from their perspective" or "when I am reading an interesting story or novel I imagine how I would feel if the events in the story were happening to me"). Questions assessing emotional empathy, on the other hand, measure the feelings of compassion and concern for others (*e.g.*, "I often have tender, concerned feelings for people less fortunate than me"). Tendency for emotional distress is also considered as a part of emotional empathy that is tapped with questions such as "being in a tense emotional situation scares me"(Davis, 1983).

Studies conducted in patients with focal brain damage suggest that cognitive and emotional empathy are dissociable at the level of brain functional organization. When comparing cognitive and emotional empathy scores in brain-damaged patients, it has been observed that damage to ventromedial prefrontal cortical structures are associated with lack of cognitive empathy. Damage to inferior frontal gyrus, a brain structure also implicated in mirroring of others' behavior, was associated with low scores on emotional empathy (Shamay-Tsoory et al., 2009). These lesion sites are illustrated in **Figure 11-5** below.

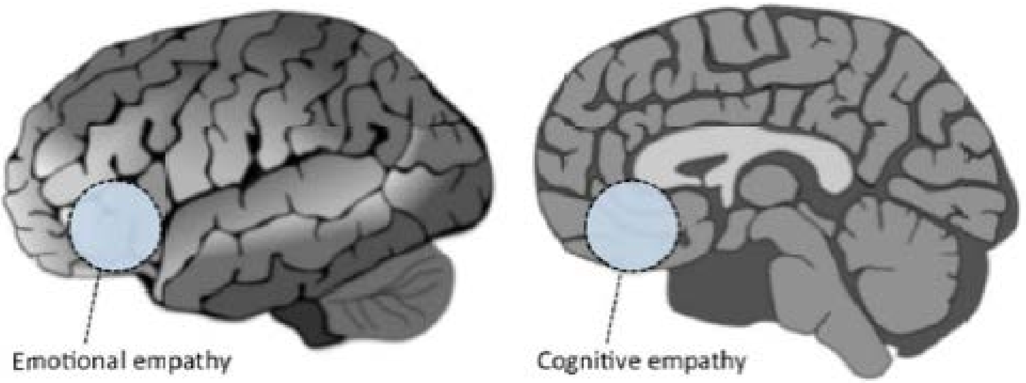


Figure 11-5. Schematic illustration showing loci of brain lesions in associated with impairments in (left) emotional empathy, and (right) cognitive empathy. Patients with emotional empathy deficits exhibit lesions in the area of inferior frontal gyrus, whereas patients with reduced cognitive empathy have lesions in ventromedial prefrontal cortex. After (Shamay-Tsoory, et al. 2009).

Indirect measures of emotions

Sometimes there are doubts whether experimental subjects want to be honest in self-reporting their emotions, even though this is typically not a concern, as subjects tend to feel obligated to help the experimenter and respond truthfully (for a famous example of this, see (Milgram, 1974)). One behavioral method that has been used is based on measuring reaction time differences to emotional and non-emotional words (Graves et al., 1981). Video recording of subjects and assessment of emotional responses based on expressions and non-verbal behavior is another example of indirect (*i.e.*, non-self report) measure of emotions (Schachter and Singer, 1962).

The video recording method has been used especially in social psychology experiments where the interactions between two or more individuals are the question of interest and, for instance, in certain clinical psychiatry studies where emotional interactions between the patient and their family members are of interest. Methods have been developed to assess the emotional expressions based on pictures or video records, such as the facial action coding system developed by Ekman and his colleagues (Ekman et al., 1978), where proficient coders track the action units (*i.e.*, specific muscle groups) of the face that can be related to specific emotional states and their intensity. Observing approach-withdrawal behaviors or fight-or-flight responses in experimental animals constitutes another set of methods that are used to indirectly assess emotions.

Measuring autonomic nervous system reactions to emotional stimuli

Measurement of physiological reactions to emotional stimuli and situations constitute another set of indirect methods that can be used to assess intensity of emotions in experimental subjects. For example, measuring heart rate changes has shown accelerated heart rate during induction of both negative and positive mental imagery (Sinha et al., 1992). On the other hand, when presented with emotional stimuli, what is typically observed is deceleration of heart rate, especially when the stimuli elicit negative emotions (Greenwald et al., 1989). Heart rate changes are regulated by the balance between sympathetic and parasympathetic systems: the parasympathetic system reduces arousal whereas the sympathetic system pushes up level of arousal, including heart rate, in preparing the body to react in an emergency situation (*i.e.*, the fight or flight response). In addition to heart rate changes, the balance between these two systems regulates blood pressure and other autonomic nervous system reactions.

Measurement of autonomic nervous system reactions, such as galvanic skin conductance responses, has been utilized in so-called lie detectors (that is also known as the polygraph). In a typical polygraph session, the suspect is presented different pieces of information, for instance, neutral items and items connected to a crime. It is then surmised that a guilty person with detailed knowledge of the crime scene will exhibit elevated autonomic nervous system activity when presented with the items connected with the crime, as compared with responses to the neutral items (Ben-Shakhar and Elaad, 2003).

Activity of facial muscle groups reveals specific emotions

The shortcoming of the autonomic nervous system measures is that whereas the level of arousal can be assessed, it is difficult to assess whether the subjects are experiencing negative or positive emotions based on arousal changes alone. Measurement of electrical activity of different facial muscles with electromyography (EMG) is a method that can be used

to assess positive *vs*. negative emotional reactions of experimental subjects (Lang et al., 1993). The activity of the muscle that draws eyebrows together (corrugator supercilii) is associated with negative emotions. The muscle associated with smiling, zygomaticus major, as well as periocular muscle (orbicularis oculi) are associated with pleasant emotions. Further increasing the potential utility of EMG recordings as measures of emotions, these increases in facial EMG activity can be observed also when there are no visible facial muscle contractions / emotional expressions (Cacioppo et al., 1986).

Use of neuroimaging methods to assess emotions

Even though neuroimaging methods have been primarily used for mapping the neural basis of emotions, rapidly accumulating knowledge of specific brain regions that exhibit enhanced activity during different emotional states is also increasingly being used to assess emotions experienced by human subjects. Neuroimaging protocols that are used to detect deception are one example of this. It seems that especially prefrontal cortex activity seems to be elevated during deception (Abe, 2009). The measurement of autonomic nervous system, facial muscle activity and brain responses provides several objective measures of emotions. However, given that there is always noise in addition to signal in physiological recordings/neuroimaging and the fact that such measures are more ambiguous to interpret than answers to direct questions, self-reports of emotional states are still widely used and valid methods for assessing emotions in experimental subjects.

11.5 Neural basis of emotions

Until the advent of modern non-invasive neuroimaging methods, observation of emotional deficits in patients with focal brain damage was the method that produced, along with animal studies, most of what is known about the neural basis of emotions. During the last two decades neuroimaging methods have enabled detailed investigation of the neural basis of emotions. So far in the present chapter, there have already been hints on the neural basis of emotions; brain structures such as amygdala, insula, and orbitofrontal cortex have been introduced as emotional areas of the brain. In the following, the specific roles that these (and certain other) brain areas play in emotions will be described.

Hemispheric differences in processing of emotions

Differences between the effects of left and right hemisphere damage on emotions constitute one of the highly consistent observations in brain-damaged patients. Right-hemisphere lesions, especially to the inferior-lateral-anterior parts of the brain, can result in patients feeling euphoric despite their plight (*e.g.*, often being crippled with paralysis of the left side) (Starkstein et al., 1989). Conversely, patients with anterior left hemisphere damage tend to feel depressed (Morris et al., 1996). In neuroimaging studies, it has been observed that there is stronger right hemisphere activity during processing of negative stimuli and stronger left hemisphere activity during processing of positive stimuli (Turhan et al., 1998). These observations are in line with the reports of brain-damaged patients; damage to the left hemisphere, for instance, results in the preserved right hemisphere activity dominating emotional processing and thus in negative mood. Damage to the right hemisphere, in turn, results in the intact left hemisphere dominating emotional processing. Since the left hemisphere is involved in processing positive moods, the patient can exhibit euphoria.

Emotions and the limbic system

Early on the so-called limbic system was implicated in emotional processing and regulation of emotions. The limbic system (see **Figure 3-10**) is a functionally heterogeneous set of brain structures located between the cerebral cortex and midbrain structures, including hippocampus, amygdala, mammillary body, cingulate and dental gyri as well as nucleus accumbens. Sometimes orbitofrontal cortex is also referred to as a limbic structure. In modern cognitive neuroscience the focus has been on delineating the functions of the distinct brain structures that are included in the limbic system rather than limbic system as a whole, which is often viewed as more of a historical concept. Amygdala, a brain structure that is located just anterior to the hippocampus in the medial temporal lobe has emerged as perhaps the most intensively studied of the limbic system structures in emotion research.

Amygdala and perception of emotional stimuli

As briefly described above, the amygdala is not a uniform structure, but is rather composed of several nuclei that are tightly interconnected (see **Figure 11-4**). The two nuclei that have received by far the most attention in cognitive neuroscience are the lateral nucleus and the central nucleus. The lateral nucleus of the amygdala is targeted by both cortical and subcortical (thalamic) inputs. The central nucleus, in turn, is the output nucleus of the amygdala that triggers hormonal, autonomic nervous system, and behavioral responses, as well as modulation of cortical processing of emotional stimuli.

The role of amygdala in processing of emotions was initiated by pioneering research where changes in emotional behavior of non-human primates were noted following medial temporal lobe lesions (Klüver and Bucy, 1939, LeDoux, 1992). For instance, it was observed that monkeys with bilateral medial temporal lobe lesions encompassing the amygdala became fearless with respect to stimuli such as snakes that would otherwise elicit strong fear responses such as freezing. In human brain-damaged patients with focal amygdala lesions, it has been subsequently observed that whereas the ability to discriminate identity information of face stimuli is preserved, the patients are unable to detect emotional expressions, especially fearful expressions, of facial stimuli. This defect was recently attributed to failure to look normally at the eye region of the face that provides the most important visual cues for recognition of fearful expressions (Adolphs et al., 2005).

In neuroimaging studies, it has been observed that amygdala is activated upon presentation of emotional stimuli in general, especially those of negative, but also ones of positive, valence. As was already shown in **Figure 11-4** above, the amygdala plays a vital role in fear conditioning. It has also been shown that the amygdala responds to fear-eliciting stimuli very quickly and even pre-attentively (Rotshtein et al., 2010). Naturally, this is highly helpful for survival of organisms, given that the organism can detect and react to potentially threatening stimuli as quickly as possible.

Recent neuroimaging studies have further shed light on how the amygdala modulates cortical processing of stimuli. It appears that stimuli with emotionally relevant information modulate cortical processing *via* ascending connections from amygdala in, for instance, the face processing areas of fusiform gyrus located in the inferior temporal-occipital cortex (Vuilleumier et al., 2004). These modulatory signals potentially enable determination of emotional facial information that takes place in the fusiform gyrus concomitantly with facial identity and gaze direction, which is a critical social signal (see Chapter 11).

Notably, the role of the amygdala is not limited to processing of emotional stimuli and guiding of behavior. The close proximity and interconnectivity between the amygdala and hippocampus explains why it is easier to remember emotional life events than emotionally neutral life events. However, highly traumatic events can be exceptions to this rule; during a very strong trauma, a person might be able to recall less well what happened. One potential explanation for this observation is that traumatic experiences elicit massive stress responses and stress hormones have been observed to reduce neurogenesis in hippocampus. For a review of findings suggesting this, see (McEwen, 1999).

Orbitofrontal cortex and evaluation of emotional cues

The orbitofrontal cortex, pictured in **Figure 3-8**, has been observed to be important for the evaluation of contextual and physiological emotional cues. This evaluation process is assumed to give rise to the cognitive percept of emotions that was discussed above when describing the study where subjects differentially interpreted their autonomic nervous system activation (that was caused unbeknownst to them by the adrenalin injection) based on whether they were in the same room with the joyful *vs.* angry person (Schachter and Singer, 1962). One of the first documented observations of the relationship between orbitofrontal cortex and goal-directed behavior was the famous patient case of Phineas Gage that was introduced in the beginning of Chapter 10. Naturally, the orbitofrontal cortex, similarly to amygdala, is not a uniform brain structure, but consists of anatomically and functionally differentiable subregions, as shown in Figure **10-2**.

According to a so-called somatic marker hypothesis, the emotional percept generated in the orbitofrontal cortex guides goal-directed behavior by providing internal feedback signals about the future consequences (*e.g.*, dangers, benefits) of behavioral actions that one is planning (Damasio, 1996). As an example of how the evaluation process influences behavior, patients with damage to orbitofrontal and ventromedial prefrontal cortical areas fail to take contextual information into account and end up behaving in socially unacceptable manner, for instance, patients with damage to these areas might take off their clothes if they feel too warm even if other people are present. Normally, one would reject the thought of taking off one's clothes to cool down as one producing feelings of embarrassment. For a review of the somatic marker hypothesis and its role in decision making, see (Bechara and Damasio, 2005).

Healthy humans are guided by positive and negative emotional signals that help choose rewarding actions and protect against getting engaged in counterproductive behaviors, such as socially inappropriate behaviors that result in social sanctions (it has to be kept in mind that social acceptance / rejection constitute one of the most powerful rewards and punishments for healthy humans). In a study where healthy controls and patients were playing a card game, both patients and controls generated skin conductance responses (a measure of autonomic system response) after selecting cards that resulted in penalties or rewards. After several trials the healthy control subjects began to generate skin conductance responses also prior to their selection of a card while they considered from which deck to choose, however, none of the patients showed these anticipatory skin conductance responses. It was concluded that the absence of anticipatory skin conductance responses in these patients is a correlate of their insensitivity to future outcomes (Bechara et al., 1996).

The role of anterior prefrontal cortex areas in assessing future outcomes of one's behavioral choices is also reflected in that these brain areas participate in risk taking behavior. This has been investigated by giving experimental subjects monetary rewards (*i.e.*, small sums of money that accumulate throughout the experiment) and punishments (*i.e.*, loosing some of the money that one has accumulated during the experiment). By manipulating the pay-off, stakes, and the probability of winning/loosing, situations can be created where the subjects engage in risk taking behaviors. It has been reported in recent neuroimaging studies that the medial/anterior orbitofrontal cortex is robustly activated both prior to accepting high value gambles and rejecting low value gambles (Cunningham et al., 2009).

Sympathetic and parasympathetic systems and emotional intensity

The sympathetic and parasympathetic systems of the autonomic nervous system are important mechanisms that regulate the intensity of experienced emotions. Patients with disconnection of ascending feedback connections from the periphery due to cervical damage have been reported to experience emotional valence, but in a less arousing manner than prior to the accident (Hohmann, 1966). The relationship between intensity of experienced emotions and autonomic nervous system activity is also targeted by specific drugs, beta-blockers, that alleviate anxiety in, for instance, performers suffering from stage fright (Brantigan et al., 1982). The beta-blockers antagonize peripheral β-adrenergic receptors, preventing increases in heart rate and other autonomic nervous system activity due to acute stress. In addition to the peripheral β-adrenergic receptors specific neurotransmitter systems in the brain also play an important role in the processing of emotional stimuli.

Neurochemistry of emotion and motivation

In addition to the role of dopamine in processing of rewards and motivation that was described earlier in this chapter, exogenous and endogenous agonists of the opioid system such as morphine or beta-endorphin, respectively, cause intense pleasure / euphoria. Conversely, blocking the brain opioid receptors with opioid antagonists such as naloxone, naltrexone, or nalmefene result in feelings of dysphoria (Mendelson et al., 1978). This effect has been used in recently developed treatments of alcoholism where the rewarding effects of alcohol drinking are blocked with simultaneous administration of opioid antagonists, which results in reduced craving for alcohol in alcohol dependent patients thus helping them remain abstinent (Sinclair, 2001).

The serotonin system also plays a role in regulation of emotions, especially in the regulation of mood. It has been described that depressed patients have reduced serotonin activity and drugs that elevate the central levels of serotonin, selective serotonin reuptake inhibitors, alleviate the symptoms of depression (Benfield et al., 1986). What makes it difficult to pinpoint the specific roles of any single neurotransmitter in complex neurocognitive phenomena such as emotions is that each of the neurotransmitters is tightly coupled with the other neurotransmitters. Thus, changes in the levels of one neurotransmitter may exert their effects *via* influencing other neurotransmitter systems. This is a potential explanation for the puzzling phenomenon of selective serotonin reuptake inhibitors alleviating depression typically with a delay of a couple of weeks, even though the levels of serotonin in the synaptic cleft are elevated in a matter of hours after administration of the drug.

11.6 Concluding comments

Emotions are highly intertwined with perceptual and cognitive processes and constitute the motivational foundation for human goal-directed behavior. Despite the paramount significance of emotions, the underlying neural mechanisms have not been investigated as extensively as those of other perceptual and cognitive processes until quite recently. Partly this has been due to limitations in neuroimaging methods in the past. While there are still multiple open questions, a pattern of results is emerging where amygdala seems to play a crucial role in fast detection of emotionally evocative stimuli, and orbitofrontal / ventromedial prefrontal cortical structures make it possible to interpret contextual and physiological cues to experience specific emotions that then guide behavioral choices. Several other brain structures participate in emotional processing as well, for instance, insula that is often associated with feelings of disgust, is thought to convey visceral information during emotion processing together with somatosensory cortex. Emotions are especially important for social cognition that will be described in the next chapter.

11.7 References

ABE, N. 2009. The neurobiology of deception: evidence from neuroimaging and loss-of-function studies. *Curr Opin Neurol,* 22**,** 594-600.

ADOLPHS, R., GOSSELIN, F., BUCHANAN, T. W., TRANEL, D., SCHYNS, P. & DAMASIO, A. R. 2005. A mechanism for impaired fear recognition after amygdala damage. *Nature,* 433**,** 68-72.

BARRETT, L. F., MESQUITA, B., OCHSNER, K. N. & GROSS, J. J. 2007. The experience of emotion. *Annual Review of Psychology,* 58**,** 373-403.

BECHARA, A. & DAMASIO, A. R. 2005. The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior,* 52**,** 336-372.

BECHARA, A., TRANEL, D., DAMASIO, H. & DAMASIO, A. R. 1996. Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cerebral Cortex,* 6**,** 215-225.

BEN-SHAKHAR, G. & ELAAD, E. 2003. The validitly of psychophysiological detection of information with the guilty knowledge test: a meta-analytic review. *Journal of Applied Psychology,* 88**,** 131-151.

BENFIELD, P., HEEL, R. C. & LEWIS, S. P. 1986. Fluoxetine. A review of its pharmacodynamic and pharmacokinetic properties, and therapeutic efficacy in depressive illness. *Drugs,* 32**,** 481-508.

BERRIDGE, K. C. & ROBINSON, T. E. 1998. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Reviews,* 28**,** 309-369.

BRANTIGAN, C. O., BRANTIGAN, T. A. & JOSEPH, N. 1982. Effect of beta blockade and beta stimulation on stage fright. *American Journal of Medicine,* 72**,** 88-94.

CACIOPPO, J. T. & BERNTSON, G. G. 1994. Relationships between attitudes and evaluative space: a critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin,* 115**,** 401-423.

CACIOPPO, J. T., PETTY, R. E., LOSCH, M. E. & KIM, H. S. 1986. Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology,* 50**,** 260-268.

CANNON, W. B. 1929. *Bodily changes in pain, hunger, fear and rage,* New York, Appleton.

CUNNINGHAM, W. A., KESEK, A. & MOWRER, S. M. 2009. Distinct orbitofrontal regions encode stimulus and choice valuation. *Journal of Cognitive Neuroscience,* 21**,** 1956-1966.

DAMASIO, A. R. 1996. The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philos Trans R Soc Lond B,* 351**,** 1413-1420.

DAVIS, M. H. 1983. Measuring individual differences in empathy: evidence for a multidimensional approach. *J Pers Soc Psychol,* 44**,** 113-126.

EKMAN, P. 1999. Facial expressions. *In:* DALGLEISH, T. & POWER, M. (eds.) *Handbook of Cognition and Emotion.* New York: John Wiley & Sons.

EKMAN, P., FRIESEN, W. V. & HAGER, J. 1978. *Facial action coding system,* Palo Alto, CA, Consulting Psychologist Press.

EKMAN, P. & HEIDER, K. G. 1988. The universality of contempt expression: a replication. *Motivation and Emotion,* 12**,** 303-308.

EKMAN, P., SORENSON, E. R. & FRIESEN, W. V. 1969. Pan-cultural elements in facial displays of emotion. *Science,* 164**,** 86-88.

FUSAR-POLI, P., PLACENTINO, A., CARLETTI, F., LANDI, P., ALLEN, P., SURGULADZE, S., BENEDETTI, F., ABBAMONTE, M., GASPAROTTI, R., BARALE, F., PEREZ, J., MCGUIRE, P. & POLITI, P. 2009. Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *J Psychiatry Neurosci,* 34**,** 418-432.

GRAVES, R., LANDIS, T. & GOODGLASS, H. 1981. Laterality and sex differences for visual recognition of emotional and non-emotional words. *Neuropsychologia,* 19**,** 95-102.

GREENWALD, M. K., COOK, E. W. & LANG, P. J. 1989. Affective judgement and psychophysiological response: Dimensional covariation in the evaluation of pictorial stimuli. *Journal of Psychophysiology,* 3**,** 51-64.

HOHMANN, G. W. 1966. Some effects of spinal cord lesions on experienced emotional feelings. *Psychophysiology,* 3**,** 526-534.

JAASKELAINEN, I. P., KOSKENTALO, K., BALK, M. H., AUTTI, T., KAURAMÄKI, J., POMREN, C. & SAMS, M. 2008. Inter-subject synchronization of prefrontal cortex hemodynamic activity during natural viewing. *The Open Neuroimaging Journal,* 2**,** 14-19.

JAMES, W. 1890. *The principles of psychology,* New York, Henry Holt.

KLÜVER, H. & BUCY, P. C. 1939. Preliminary analysis of functions of the temporal lobes in monkeys. *Archives of Neurology and Psychiatry,* 42**,** 979-1000.

LANG, P. J., BRADLEY, M. M. & CUTHBERT, B. N. 2005. *International affective picture system (IAPS): affective ratings of pictures and instruction manual,* Gainesville, FL, University of Florida.

LANG, P. J., GREENWALD, M. K., BRADLEY, M. M. & HAMM, A. O. 1993. Looking at pictures: affective, facial, visceral, and behavioral reactions. *Psychophysiology,* 30**,** 261-273.

LEDOUX, J. E. 1992. Emotion and the amygdala. *In:* AGGLETON, J. P. (ed.) *The amygdala: neurobiological aspects of emotion, memory, and mental dysfunction.* New York: Wiley-Liss.

MCEWEN, B. S. 1999. Stress and hippocampal plasticity. *Annual Review of Neuroscience,* 22**,** 105-122.

MENDELSON, J. H., ELLINGBOE, J., KEUHNIE, J. C. & MELLO, N. K. 1978. Effects of naltrexone on mood and neuroendocrine function in normal adult males. *Psychoneuroendocrinology,* 3, 231-236.

MILGRAM, S. 1974. *Obedience to authority,* New York, Harper and Row.

MORRIS, P. L., ROBISON, R. G., RAPHAEL, B. & HOPWOOD, M. J. 1996. Lesion location and poststroke depression. *Journal of Neuropsychiatry and Clinical Neurosciences,* 8, 399-403.

ONTON, J. & MAKEIG, S. 2009. High-frequency broadband modulations of electroencephalographic spectra. *Front Hum Neurosci,* 3, 61.

PATON, J. J., BELOVA, M. A., MORRISON, S. E. & SALZMAN, C. D. 2006. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature,* 439, 865-870.

PECINA, S. & BERRIDGE, K. C. 2000. Opioid site in nucleus accumbens shell mediates eating and hedonic 'liking' for food: map based on microinjection Fos plumes. *Brain Research,* 863, 71-86.

ROGAN, M. T., STÄUBLI, U. V. & LEDOUX, J. E. 1997. Fear conditioning induces associative long-term potentiation in the amygdala. *Nature,* 390, 604-607.

ROTSHTEIN, P., RICHARDSON, M. P., WINSTON, J. S., KIEBEL, S. J., VUILLEUMIER, P., EIMER, M., DRIVER, J. & DOLAN, J. R. 2010. Amygdala damage affects event-related potentials for fearful faces at specific time windows. *Human Brain Mapping,* 31, 1089-1105.

RUSSELL, J. A. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology,* 39, 1161-1178.

RUSSELL, J. A. & BARRETT, L. F. 1999. Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *Journal of Personality and Social Psychology,* 76, 805-819.

SCHACHTER, S. & SINGER, J. 1962. Cognitive, social and physiological determinants of emotional state. *Psychological Review,* 69, 379-399.

SCHLOSBERG, H. 1954. Three dimensions of emotion. *Psychological Review,* 61, 81-88.

SCHNEIDER, F., GRODD, W., WEISS, U., KLOSE, U., MAYER, K. R., NÄGALE, T. & GUR, R. C. 1997. Functional MRI reveals left amygdala activation during emotion. *Psychiatry Research: Neuroimaging,* 76, 75-82.

SCHULZ, W. 2000. Multiple reward signals in the brain. *Nature Reviews Neuroscience,* 1, 199-207.

SHAMAY-TSOORY, S. G., AHARON-PERETZ, J. & PERRY, D. 2009. Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain,* 132, 617-627.

SINCLAIR, J. D. 2001. Evidence about the use of naltrexone and for different ways of using it in the treatment of alcoholism. *Alcohol Alcohol,* 36, 2-10.

SINHA, R., LOVALLO, W. R. & PARSONS, O. A. 1992. Cardiovascular differentiation of emotions. *Psychosomatic Medicine,* 54, 422-435.

STARKSTEIN, S. E., ROBINSON, R. G., HONIG, M. A., PARIKH, R. M., JOSELYN, J. & PRICE, T. R. 1989. Mood changes after right-hemisphere lesions. *British Journal of Psychiatry,* 155, 79-85.

TURHAN, C., DESMOND, J. E., ZHAO, Z., GLOVER, G. & GABRIELI, J. D. E. 1998. Hemispheric asymmetry for emotional stimuli detected with fMRI. *NeuroReport,* 9, 3233-3239.

VIINIKAINEN, M., JÄÄSKELÄINEN, I. P., ALEXANDROV, Y., BALK, M. H., AUTTI, T. & SAMS, M. 2010. Nonlinear relationship between emotional valence and brain activity: evidence of separate negative and positive valence dimensions. *Human Brain Mapping,* 31, 1030-1040.

VUILLEUMIER, P., RICHARDSON, M. P., ARMONY, J. L., DRIVER, J. & DOLAN, J. R. 2004. Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nature Neuroscience,* 7, 1271-1278.

WESTERMANN, R., SPIES, K., STAHL, G. & HESSE, F. W. 1996. Relative effectiveness and validity of mood induction procedures: a meta-analysis. *European Journal of Social Psychology,* 26, 557-580.

ZENTNER, M., GRANDJEAN, D. & SCHERER, K. R. 2008. Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion,* 8, 494-521.

12 Social cognition

If one keeps a mental log of things that are on one's mind during a typical day, one soon realizes that it is the social relationships with other people, at work, home, hobbies, that occupy one's thoughts the vast majority of time. On the other hand, if one meets someone one has not met in several years, one will recall with relative ease the relevant aspects of his/her personality and past interactions with him/her, allowing one to quickly "tune in" to interact with him/her. Observations such as these suggest that processing social aspects of the environment, such as perceiving gestures/expressions and pondering the intentions of others, is a major task that the human brain is constantly engaged in. The perceptual and cognitive functions that support this are collectively called social cognition.

Indeed, the human species is inherently social and most of what humans have accomplished is based on working in groups, from hunting in groups in the prehistoric times to accomplishing modern technological advances in teams of experts today. Notably, in such teamwork social interactions do enhance "regular cognition" as well; it is often through exchange of knowledge with others that individuals of a team can solve problems and come up with novel insights. As yet another instance of the social nature of human species, social isolation is perhaps the harshest of punishments that are used in prisons. Much like the research on the neural basis of emotions, however, studies on the neural underpinnings of social cognition have not become a topic of intensive research until relatively recently. Recent findings from evolution studies have especially stimulated research into the neural basis of social cognition.

12.1 Demands on social cognition predict evolution of cortex

The notion that social cognition is a central task for the brain is supported by findings in evolutionary studies where the number and depth of social interactions—and especially more intensive forms of pair-bonding—that a given species has to manage is the factor that best explains the size of the neocortex (Dunbar and Shultz, 2007). The increase in the ratio of neurocortex to the rest of the brain as a function of mean social group size in non-human primates is shown in **Figure 12-1**. Further, there is evidence suggesting that the growth of especially the most anterior part of the frontal cortex (frontal pole / area 10) during phylogeny is best explained by the social cognition demands that humans face (Semendeferi et al., 2001).

These evolution studies, constituting the foundation for the so-called "social brain hypothesis", have contrasted the previously held belief that it would be the classical intelligence (as measured by the intelligence quotient in humans) that explains why humans have larger neocortex than other animals. Naturally, there are highly developed cognitive abilities in humans that support also other functions/behaviors besides social cognition. One example of such functions is the ability of humans to delay gratification, meaning that one can pursue rewards/goals that are quite distant in time and choose not to take alternative rewards that are immediately accessible. This ability is disturbed following brain damage to the ventral-anterior parts of the frontal lobes, which is also important for social cognition.

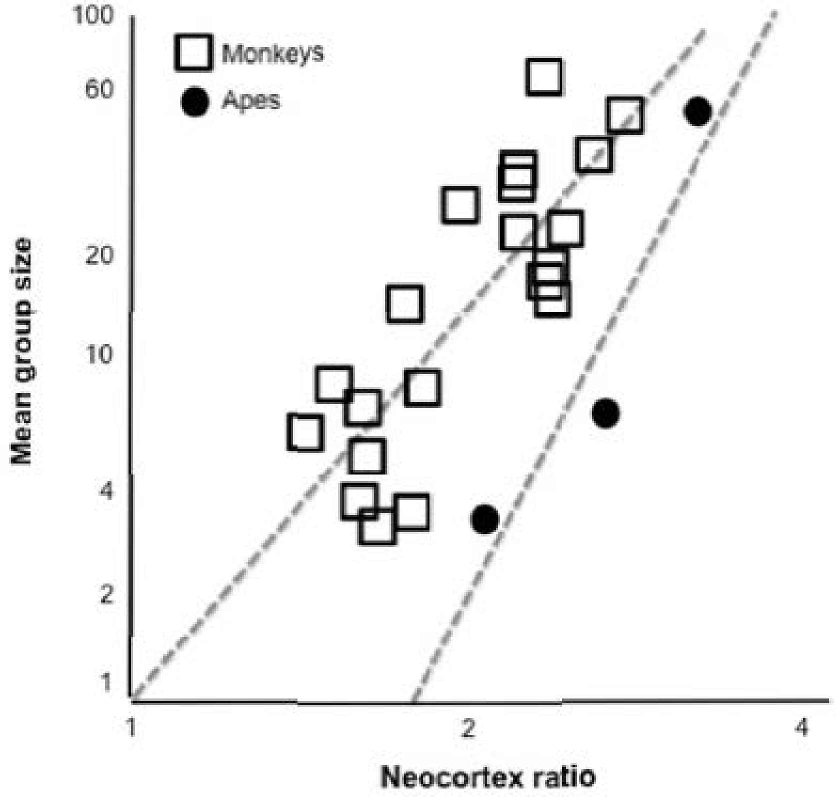


Figure 12-1. Shown is the increase in the neocortex ratio as a function of mean social group size in non-human primates. Based on data presented in (Barrett et al., 2002).

12.2 Social cognition is composed of multiple functions

While the evolutionary studies provide robust evidence for the importance of social interactions as a factor explaining growth of neocortex, they do not directly answer the question of what type of skills it is that make smooth and effortless social interactions possible for humans. To answer this question there are, first of all, specific brain mechanisms that make it possible to detect fine-grained social cues such as facial expressions or gaze direction of others. On the other hand, the anterior brain regions that have been shaped the most in the course of evolution into social species make it possible to infer the intentions of others. In the following, the most central social cognitive functions will be described, as well as what is currently known about the underlying neural mechanisms.

Fusiform gyrus is devoted to processing of facial information

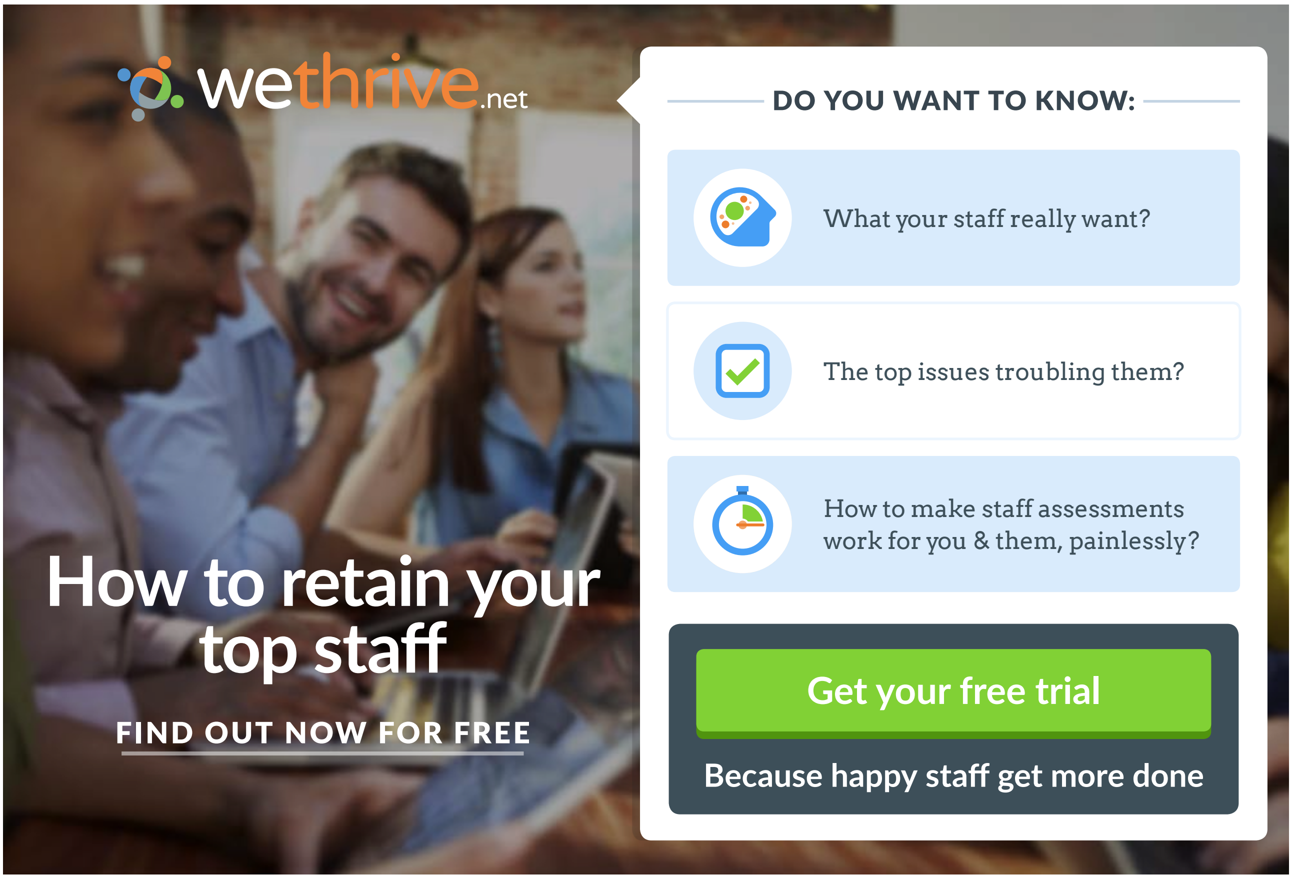
The fusiform gyrus that is located bilaterally in the inferior aspects of each hemisphere at the junction between occipital and temporal lobes has been shown to exhibit face-specific responses in both in non-human primate single-unit recordings and neuroimaging studies in humans. It has been reported that the fusiform face area exhibits larger responses to faces than to any other visual object category that has been tested (Kanwisher et al., 1997). Face-specific responses have been observed in the fusiform face area also in studies using highly naturalistic stimuli such as movies (Hasson et al., 2004, Kauppi et al., 2011).

Importantly, the fact that there is an entire brain area that is devoted to processing of facial stimuli speaks for the significance of face processing to human social cognition. In fact, already newborn infants pay close attention to facial stimuli (Goren et al., 1975). In addition to fusiform face area there are, however, neurons responsive to facial stimuli in other brain areas such as the amygdala and superior temporal sulcus (Haxby et al., 2000). As will be described next, the amygdala is especially devoted to fast processing of facial information that is vital for detecting emotionally evoking (and thus potentially important) stimuli.

Amygdala enables fast perception of emotional stimuli

The amygdala has been identified as a brain structure that processes facial emotional expressions even subliminally (Whalen et al., 1998). Subliminal means that the presentation of the stimulus is so brief (few tens of milliseconds) that a conscious percept does not occur. Typically, a masker stimulus of some kind immediately follows the brief presentation of the emotional stimulus to wipe out any retinal after images. It has been suggested that the ability of amygdala to react quikly to emotional stimuli is based on fast processing of low spatial frequencies of visual inputs (Vuilleumier et al., 2003, Bar et al., 2006).

The capability of the amygdala to process stimuli and assess their emotional significance very quickly has been also linked to phenomena such as prejudice (Winston et al., 2002). Here, with prejudice it is meant that one forms quick (and potentially highly erroneous) impressions of whether someone is trustworthy or likeable (or not trustworthy/likeable) that then guides subsequent perception of him/her. Indeed, social perception occurs iteratively where information is collected in cycles to get a more accurate picture of the disposition and intentions of others.



The iterative nature of social peception is exemplified by findings where white Americans were shown faces of black persons either for very brief durations of 30 ms (that allowed only subliminal perception) *vs*. for 525 ms. During the longer exposure, the amygdala responses were reduced and frontal cortical activity was increased, which was interpreted as suggesting that implicit prejudices were suppressed by frontal cortical areas when longer exposure times were utilized (Cunningham et al., 2004). In addition to being able to identify individuals and infer emotional states of others based on their facial expressions (which is in by itself a rather amazing ability given how small the cues can be), the direction of others' gaze is a cue that contains a wealth of information about their intentions. Findings pertaining to this will be described next.

One perceives the gaze direction of others very accurately

The ability to quickly and accurately notice the gaze direction of others is one of the fundamental abilities that help decode intentions of others. This is a rather amazing ability given how fine a visual cue gaze direction is; even if a person is standing some distance away, one is still able to tell what that person is looking at. The orientation of the head and the body convey information that complements that given by the eyes. Furthermore, faces with gaze directed at *vs*. diverted away from experimental subjects, as shown on **Figure 12-2**, seems to constitute an especially salient social cue that modulates perception of faces, interacting especially with emotional expressions that the faces display. For instance, it has been observed that perceived self-relevance of fearful faces is higher with averted gaze signaling a nearby danger and, conversely, angry faces directly gazing at the experimental subject are judged more relevant (*i.e.*, signaling aggressiveness against the subject) than angry faces with diverted gaze (N'Diave et al., 2009).

It has also been observed that directed *vs*. averted gaze interacts with perceived attractiveness of faces. Specifically, perceived attractiveness of faces with directed gaze were higher for smiling than for neutral faces. This effect was reversed when gaze was averted away; faces with neural expressions were then rated as more attractive than smiling faces (Jones et al., 2006). A person smiling at oneself is of course a very powerful social invitation, unless there is contextual information that modulates perception of the social cue (*e.g.*, if someone is interpreted as smiling derisively at oneself due to ill personal history with that person). The sensitivity to gaze direction seems to be an innate ability as newborn infants have been reported to have the ability to detect gaze directed at oneself *vs.* averted away, thus guiding development of social cognition (Guellai and Streti, 2011).

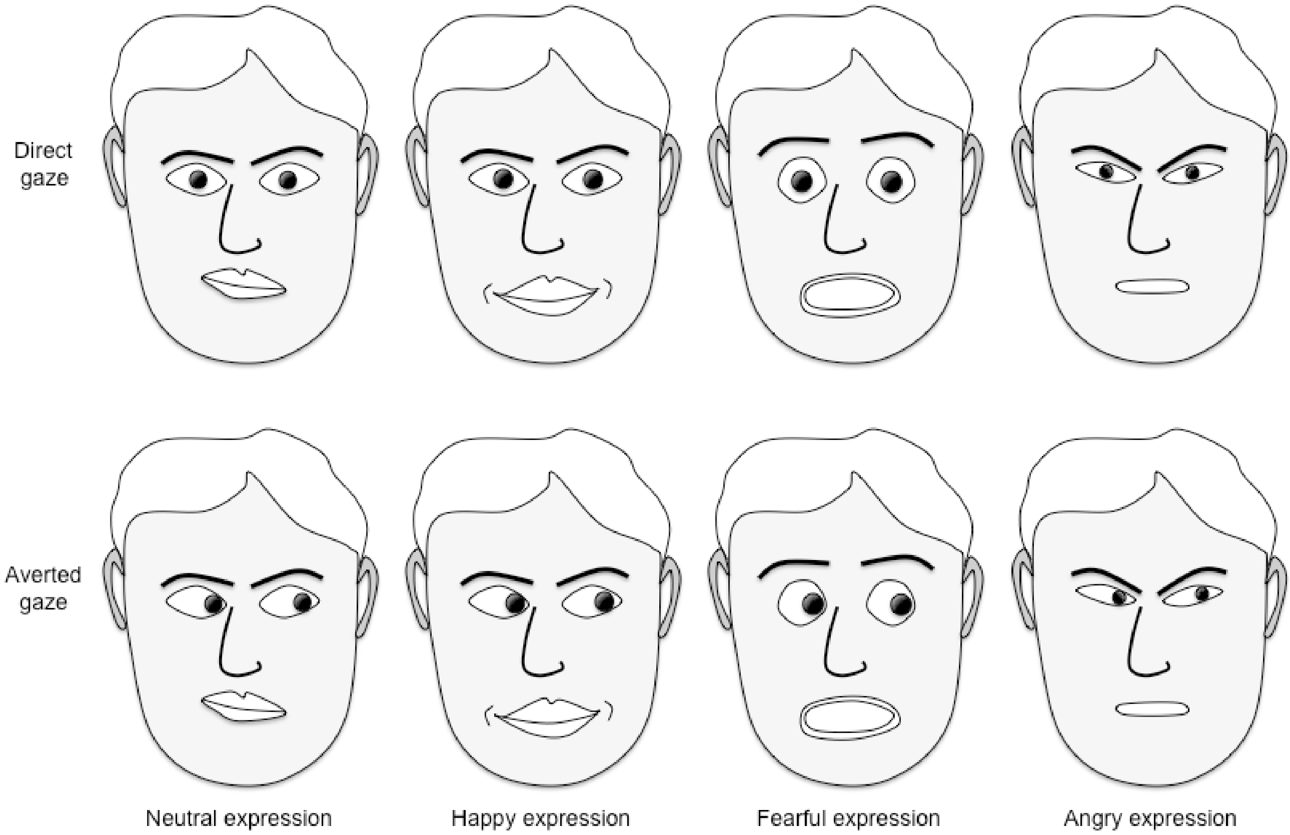


Figure 12-3. Gaze direction, when directed at oneself *vs*. averted, modulates perception of faces, especially perceived relevance and strength of emotional expressions. Here, from left to right are depicted neutral, happy, fearful and angry expressions with the only difference between the faces on the upper *vs*. lower row being that the gaze is directed *vs*. averted.

Superior temporal sulcus, amygdala and orbitofrontal cortex are important for perception of gaze direction

Neurons that are responsive to facial stimuli in non-human primate amygdala have also been reported to respond more vigorously when gaze is directed at the monkey *vs*. when averted away (Tazumi et al., 2010). While this finding could be partially explained by the fact that directed gaze is an especially powerful social cue monkeys that establishes social hierarchy among monkeys (*i.e.*, direct eye contact with the dominant male of the monkey colony would constitute a challenge), the presence of similar gaze-direction sensitive cells have been observed in human neuroimaging studies (N'Diave et al., 2009). In other studies, ablation of the superior temporal sulcus has been observed to result in failure to detect gaze direction (Campbell et al., 1990) and, additionally, the orbitofrontal cortex has been suggested to participate in detecting gaze together with the amygdala and the superior temporal sulcus (Emery, 2000). The brain areas participating in perception of gaze direction, head direction, body orientation, faces, and facial expressions are summarized in **Figure 12-3**.

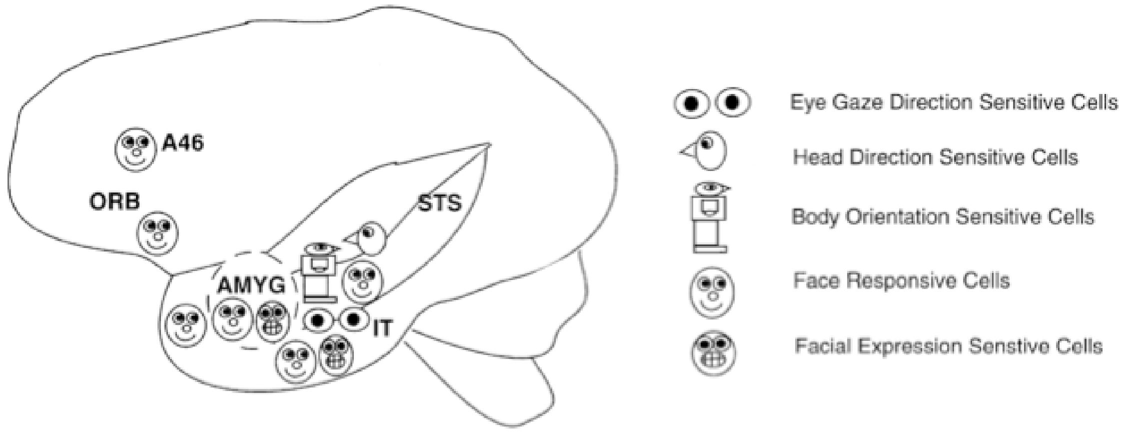


Figure 12-3. A schematic illustration summarizing brain areas that have been found sensitive to eye gaze direction, head direction, body orientation, faces, and facial expressions overlaid on the left hemisphere of a non-human primate brain. Adapted from (Emery, 2000).

Studies conducted on so-called split-brain patients have suggested that the right hemisphere is specifically relevant for detecting gaze direction of others. Split-brain patients are individuals who have had their corpus callosum connections cut during neurosurgery for the purpose of controling spreading of medication-resistant epilepsy from one hemisphere to the other. The disconnection of the inter-hemispheric connections in these patients has, however, resulted in a number of unexpected cognitive deficits that have been investigated in detail. Studies in these patients have suggested that the right hemisphere is responsible for automatic orienting of attention to gaze direction (Kingstone et al., 2000).

Gaze regulates level of intimacy in two-person interactions

Studies conducted in the field of social psychology have shown that the level of intimacy in two-person interactions is regulated by a number of factors (Argyle and Dean, 1965). The length of eye contacts varies as a function of how well one knows (and likes) the other person, and is affected by the physical distance, and depth of conversation that the two persons are engaged in. Two persons engaged in discussion on a superficial topic have longer eye contacts than ones engaged in discussing more personally sensitive issues. Likewise, being physically close to another person leads to shorter and fewer direct eye contacts with the conversation partner. Romantically involved persons of course violate these rules and can be viewed as a case where two persons want to be both physically and mentally as close as possible to each other. There are also cultural differences, with the "normal" physical distance and length of eye contact varying significantly, which can lead to awkward situations when two persons from different cultures interact if they are not cognizant of cultural differences in their personal space.

Interestingly, schizophrenia patients who have severe deficits in social skills exhibit specific problems in detection of gaze direction; these patients have been reported to misperceive gaze averted from them as a gaze that has been directed at them (Hooker and Park, 2005). This type of deficit in gaze direction perception is likely to affect adversely the ability of schizophrenia patients to infer intentions of other people. Problems in detecting directed gaze may also contribute to deficits that schizophrenia patients have been reported to have in perception of facial expressions and emotional prosody of others (Edwards et al., 2002).

Amygdala plays a role in regulation of personal space

In addition to its key role in processing emotions, lesions of amygdala have been reported to lead to a lacking of personal space (Kennedy et al., 2009). A patient with lesion of the amygdala described being completely comfortable even when other persons were standing a few centimeters from her. At the same time, she had retained the concept of personal space, by being cognizant that she might be causing discomfort to other persons if she was to stand in too close a proximity. A control fMRI study in healthy subjects corroborated these findings; amygdala showed stronger activity when the subjects knew that the experimenter was standing right next to the scanner *vs.* when the experimenter was maintaining distance to the subjects when they were being scanned (Kennedy et al., 2009).

Theory of mind

Theory of mind is a central concept in social cognition research. It refers to the ability to think about mental states and thoughts of others (Premack and Woodruff, 1978). In other words, the theory of mind refers to having a theory of the mind of others; what others are and are not aware of, what others think and believe in, the goals they are pursuing, and what they are about to do in the near future. There are two views (that need not be mutually exclusive) on cognitive processes underlying theory of mind. The first is the so-called "theory theory". The theory theory posits that sets of causal laws are learned that relate external and internal states to behaviors, thus allowing one to construct theories of the mental states of others. The "simulation theory", on the other hand, postulates that mental states of others are simulated by "mirroring" them with the help of knowledge of one's own mental states (Gallase and Goldman, 1998)

An example of a test that is utilized in probing whether the theory of mind has developed in young children is illustrated in **Figure 12-4**. In this test, a cartoon is shown where a teddy bear is put in a box so that a child that is depicted sees it.

The teddy bear is then moved away from the box while the child in the cartoon is away. The child that is being tested is then asked where the child in the cartoon begins to look for the teddy bear when she comes back. Children with developed theory of mind capability will correctly answer "the box" as they realize that the child in the cartoon who was away when the teddy bear was taken out of the box and hidden elsewhere does not have knowledge of the altered whereabouts of the teddy bear. This capability develops at around age of four to five years in healthy children (Leslie, 1987).

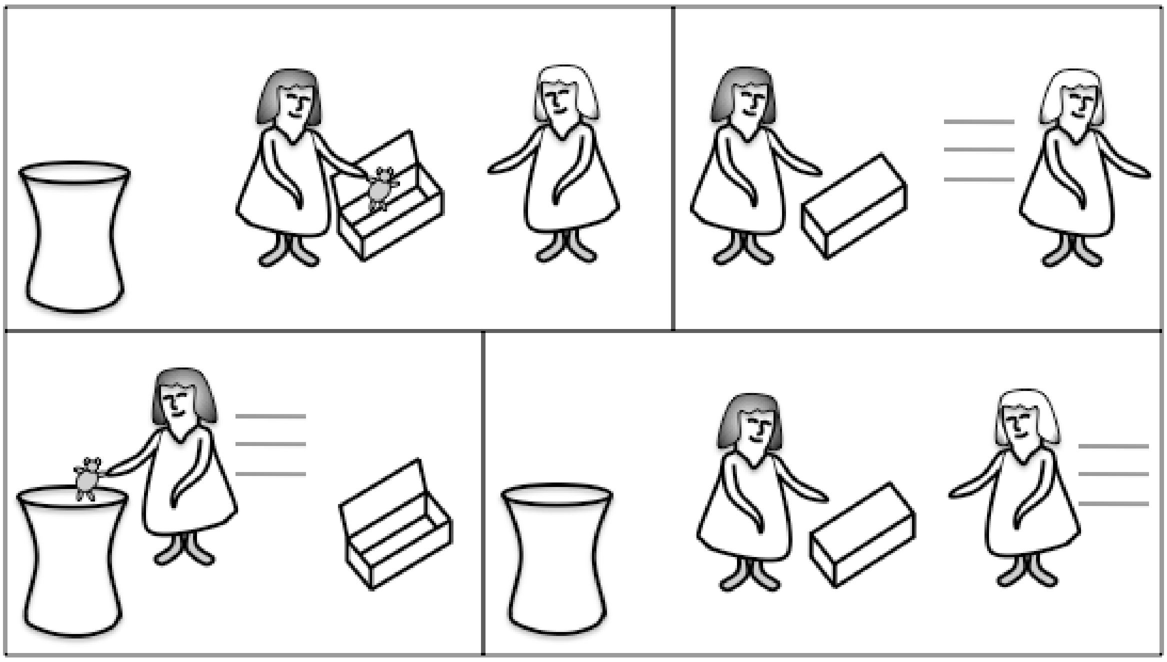


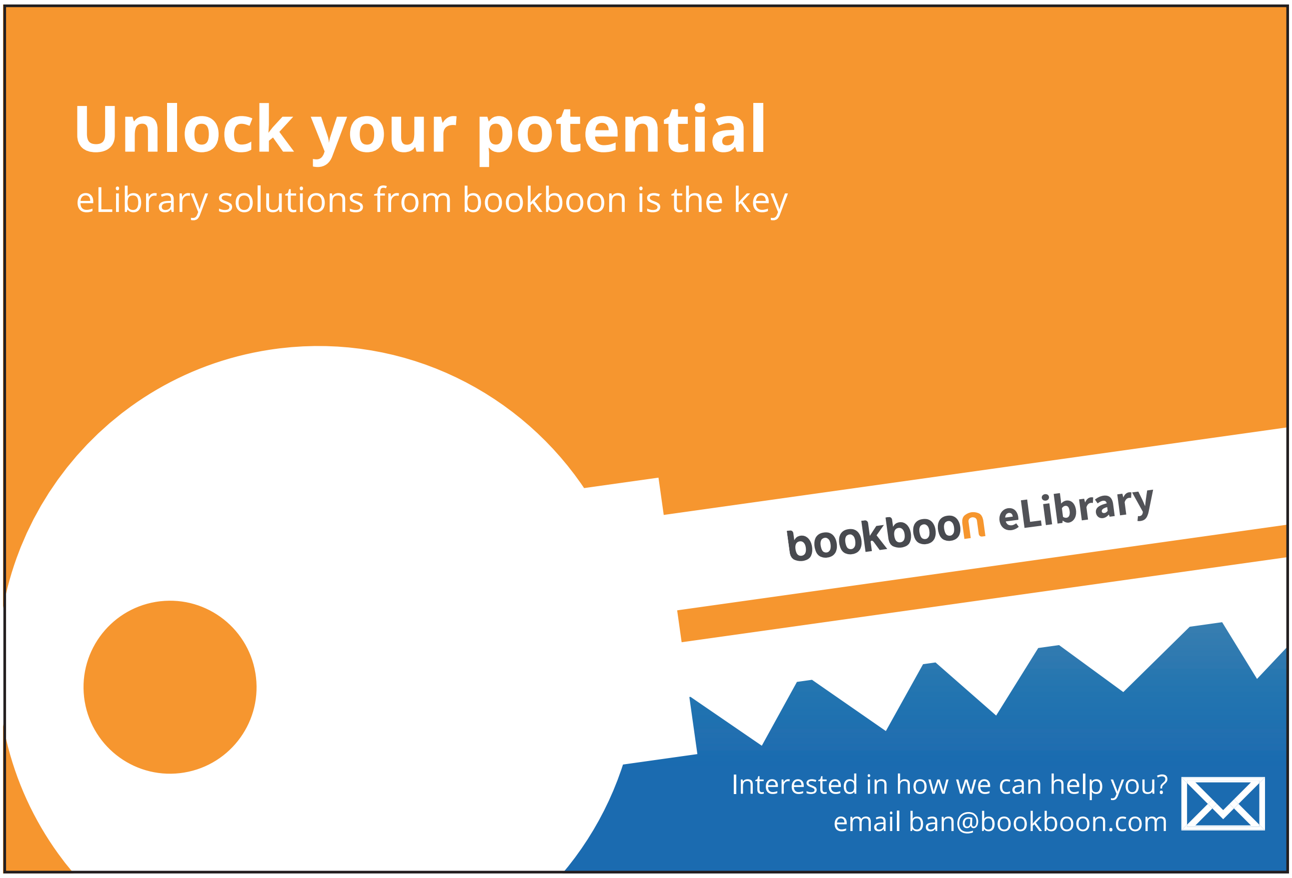
Figure 12-4: An example of the type of test that has been used to study development of the theory of mind in children. Here, a cartoon is shown to children where first a dark-haired girl puts a teddy bear into a box so that a blond-haired girl can see it. The box is then closed and the blond-haired girl exits the room. While the blond-haired girl is away, the dark-haired girl takes the teddy bear from the box and hides it into a vase. With the box again closed, the blond-haired girl re-enters the room. The child who is tested on theory of mind is then asked where the blond-haired girl will look into when searching for the teddy bear. When theory-of-mind capability is developed at around 4-5, the child correctly answers "the box" as he/she realizes that the blond girl does not have knowledge of the teddy bear being switched from the box to the vase. In contrast, younger children without theory of mind capability developed fail to grasp what the blond-haired girl in the cartoon knows and answer "the vase".

In addition to this type of cartoon tests, there are other types of theory of mind tests. For a recent review of these, see (Carrington and Bailey, 2009). For example, there are studies that have shown neurological patients cartoons or videoclips of inappropriate behaviors given the depicted social context ("faux pas") and noted deficiencies in the ability of such patients to notice the inappropriate behaviors. In non-human primate studies, theory of mind ability has been inferred based on whether a given species exhibits attempts to deceive other members of the species (*i.e.*, the so called "false belief" tests). The reasoning behind this is based on the assumption that one has to have a theory of others' minds in order to be able to infer that the others are not aware of what one is aware of (Tomasello et al., 2003).

Neural basis of the theory of mind

The neural basis of the theory of mind has been extensively studied in adults, both through documenting deficits manifested in neurological patients as a function of lesion site, and using neuroimaging methods to map brain regions specifically activated in theory of mind tasks. Prefrontal cortical regions, especially ventromedial and ortibofrontal cortical areas of the right hemisphere, seem to be vital for performance in theory of mind tasks in patients with brain damage (Martin-Rodriguez and Leon-Carrion, 2010). Findings from neuroimaging studies in healthy volunteers support these observations. In a recent meta-analysis of neuroimaging, medial prefrontal and orbitofrontal prefrontal cortical regions were the areas of the brain that were most consistently activated during a number of variations of theory of mind tests/ tasks (Carrington and Bailey, 2009).

However, there were a number of other brain areas that have been implicated in theory of mind studies as well, including lateral prefrontal cortex, supplementary motor area, motor cortex, anterior cingulate gyrus, para-cingulate gyrus, precuneous, posterior cingulate gyrus, anterior temporal lobe, superior temporal sulcus, temporo-parietal junction including intraparietal lobule, occipital cortex, insula, fusiform gyrus, and cerebellum (Carrington and Bailey, 2009). For a summary of these areas, see **Figure 12-5**. Thus, it is most likely that there is a network of brain areas that participate in theory of mind, rather than any single brain region. For instance, it is likely that the mirror neurons are utilized to mirror the emotions and actions of others when carrying out theory of mind tasks, potentially explaining activity in motor, somatosensory and insular cortex during these tasks (Frith, 2007).



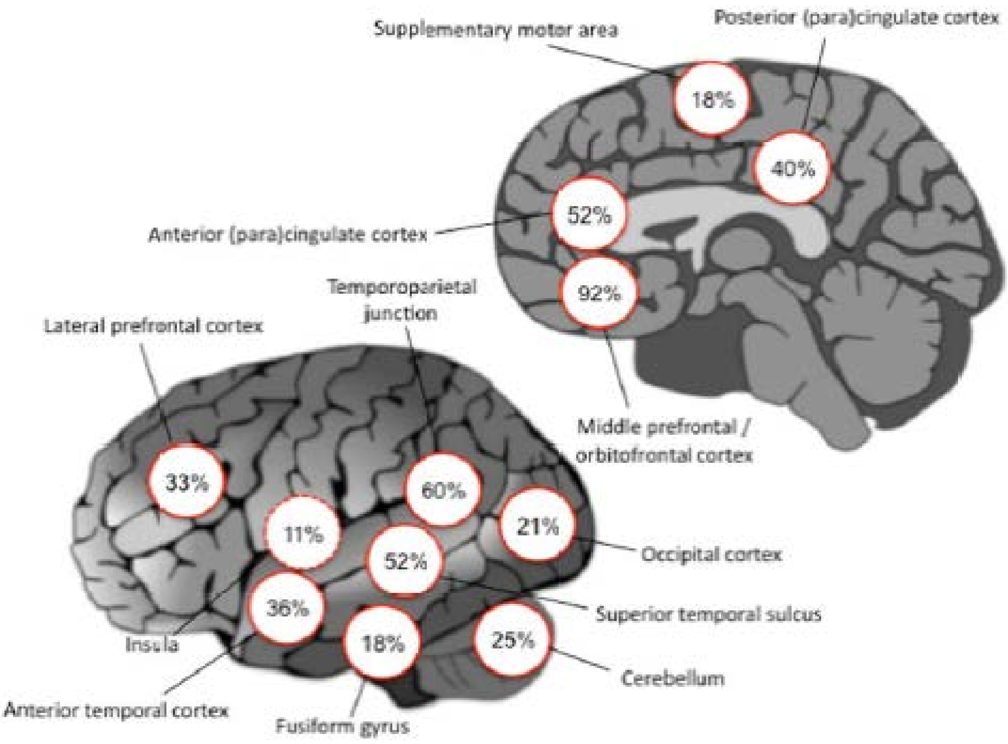


Figure 12-5. Percentage of studies in which various types of theory of mind tests activated the brain areas indicated with the circles based on a recent meta-analysis (Carrington and Bailey, 2009). As can be seen, the middle prefrontal / orbitofrontal cortical areas are most often activated during theory of mind tasks, but there is a plethora of other brain areas involved depending on the type of theory of mind task that is utilized.

Social scripts (also known as schema representations), that are formed through accumulation of information about social encounters and personalities of individuals one knows, are crucial for understanding the minds and intentions of others. Temporal poles have been implicated to play a key role in supporting such scripts (Funnell, 2001). However, it seems that the medial prefrontal cortex and superior temporal sulcus are the brain regions that are most central for theory of mind, and the other areas are then activated contingent on the specific theory of mind task requirements (Carrington and Bailey, 2009). In the following, findings linking the ventromedial prefrontal cortex with social cognition are described in more detail.

Levels of intentionality and the ventromedial prefrontal cortex

Being able to infer the intentions that other people have is an ability that is inherently related to the theory of mind. In fact, it can be argued that theory of mind can be defined as ability to decode intentionality (*e.g.*, “I think that you believe that…”). While the type of theory of mind task that is depicted in **Figure 12-4** above is very easy for any adult person, the theory of mind ability of adults can be tested by increasing the levels of complexity for to-be-inferred intentions. For instance, given the cartoon in **Figure 12-4**, first-order intentionality inference could be probed by asking what the dark-haired girl is thinking, second-order intentionality would refer to being able to infer what the dark-haired girl thought that the blond-haired girl thought, third-order intentionality what the blond-haired girl thought that the dark-haired girl thought about the blond-haired girl herself, and so forth.

It has been observed that most adults are capable of four to five levels of intentionality inference, with very rare individuals able to infer sixth order of intentionality. While sixth-order level of intentionality (or even third or fourth level) is rarely needed in everyday life, it can be argued that the relative ease at which a given person can perceive the intentions of others, and thus understand others, probably correlates with performance in the levels of intentionality test. In keeping with the central role of the ventromedial prefrontal cortex for theory of mind observed in functional neuroimaging studies, increased gray matter volume in the ventromedial prefrontal cortex has been associated with both the ability to infer levels of intentionality and size of the social network of healthy volunteers (Lewis et al., 2011).

Ventromedial prefrontal cortex processes contingent reward information

Another aspect that makes the ventromedial prefrontal cortex highly suitable brain region for social cognition is that it has been observed to be able to take higher-order contextual information and contingencies into account when learning associations between stimuli and rewards. In classical reinforcement learning studies, a given stimulus is associated with reward in a straightforward manner. There are specific brain structures, most notably striatum, that exhibit responses as the stimulus-reward associations are learned. However, the ventromedial prefrontal cortex has been shown to exhibit activity patterns that match a more complex form of learning, where reward value that is dependent on task structure can be learned (Hampton et al., 2006).

In the case of social situations requiring social cognition, of course, whether one sees a smiling person is not always reinforcing (*i.e.*, it does not always constitute a positive signal), but the reinforcing value depends on the social context. For instance, the reward value of a smile can be entirely different if a person with key talents at one's company is smiling because he/she is happy about working in one's company *vs.* if he/she is smiling because he/she has just found a better job. As another example, the reward value of a smile is highly different if the member of opposite sex that one has a crush on smiles at oneself *vs.* at some other person who is trying to get his/her attention with romantic intentions.

Going from two to three: social psychology of group dynamics

While the vast majority of neuroimaging studies of social cognition have been conducted on a single subject at a time basis (*i.e.*, showing socially relevant stimuli to a subject during neuroimaging), there are already demonstrations of neuroimaging of two persons who are interacting with one another. Interactions that occur between three or more persons have been extensively studied in a discipline called social psychology. One example of the factors that govern behavior of such groups is captured by the so-called balance theory (Heider, 1946). The balance theory argues that if positive/like and negative/dislike relations between two or more individuals are not in balance, cognitive dissonance (*i.e.*, an uncomfortable feeling that forces one to change one's attitude towards persons or things) ensues that forces the individual to change his/her sentiment towards the others.

As an example of this, in a triad consisting of three persons the balance state is, according to the balance theory, achieved if the algebraic multiplication of the signs of the triad relations is positive (see **Figure 12-6**). In the example depicted, A and B dislike one another, but A and C, as well as B and C like each other. Thus, the triad is not in balance and there should be cognitive dissonance forcing changes in the triad relationships. In this particular constellation, the triad achieves balance if 1) A and B start to like each other, or 2) if A begins to dislike C in addition to disliking B. Naturally, balance could also be obtained if C and B begin to dislike each other. Situations such as this are not difficult to envision in real life, as it is quite common that friends of friends end up becoming friends or, alternatively, if interactions with friend of a friend who one does not like are enforced, the friendship itself could be severed.

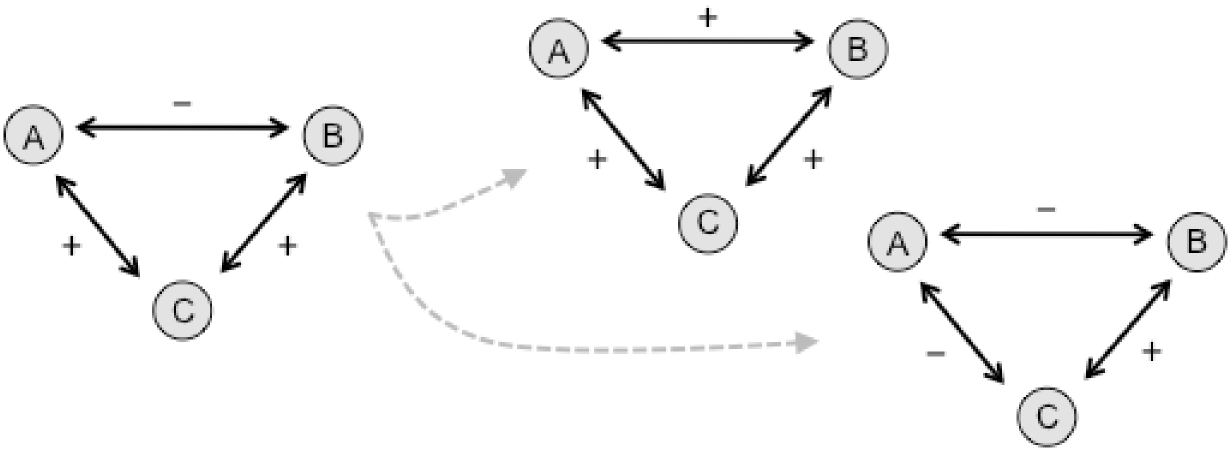


Figure 12-6. An example of how the social balance theory predicts dynamics in triad relationships. On the left, both A and B like C, but A and B do not like each other. Thus the triad is not in balance and cognitive dissonance ensues that forces changes in the relationships. This could be solved by, for instance, A and B starting to like one another, or A starting to dislike C.

While there are few neuroimaging findings that relate to group dynamics, there are observations that might be relevant. For instance, anterior cingulate cortex activation was observed in subjects when they were rejected by another person, as opposed to when they were rejected by a computer (Eisenberger et al., 2003). It is possible that the observed anterior cingulate gyrus activation is related to feeling of cognitive dissonance implied in the balance theory. However, studies on the social-neurocognitive mechanisms that govern group dynamics under real life conditions are currently scarce. This probably constitutes one of the next major research questions in the field of social cognition studies.

12.3 Concluding remarks

Social cognition is a rapidly emerging and highly exciting research area in cognitive neuroscience. Social cognition involves both perception of social and emotional cues in social contexts as well as inferring of mental states and intentions of others. Perceiving the gaze direction of others constitutes one of the most central social cues; directed *vs*. averted gaze modulates both perception of other's emotional expressions and the attractiveness of smiling faces. Gaze perception has been found to rely on a network of brain areas consisting of the superior temporal sulcus, amygdala, and orbitofrontal cortex.

One of the most central concepts in social cognition is the theory of mind, which means that people have the ability to infer the mental states and intentions of others. There is accumulating information indicating that medial prefrontal cortical areas (encompassing the orbitofrontal and ventromedial prefrontal areas) are most central for theory of mind, even though there are a number of other brain areas that network with the middle prefrontal cortex depending on the specific type of theory of mind task. Inferring about the mental states of others is also closely linked to self-awareness / consciousness. Cognitive neuroscience findings pertaining to consciousness will be described in the next chapter.

12.4 References

ARGYLE, M. & DEAN, J. 1965. Eye-contact, distance and affiliation. *Sociometry,* 28**,** 289-304.

BAR, M., NETA, M. & LINZ, H. 2006. Very first impressions. *Emotion,* 6**,** 269-278.

BARRETT, L., LYCETT, J., DUNBAR, R. 2002. *Human Evolutionary Psychology*, Basingstoke, Palgrave-Macmillan.

CAMPBELL, R., HEYWOOD, C. A., COWEY, A., REGARD, M. & LANDIS, T. 1990. Sensitivity to eye gaze in prosopagnosis patients and monkeys with superior temporal sulcus ablation. *Neuropsychologia,* 28**,** 1123-1142.

CARRINGTON, S. J. & BAILEY, A. J. 2009. Are there theory of mind regions in the brain? A review of the neuroimaging literature. *Human Brain Mapping,* 30**,** 2313-2335.

CUNNINGHAM, W. A., JOHNSON, M. K., RAYE, C. L., GATENBY, J. C., GORE, J. C. & BANAJI, M. R. 2004. Separable neural components in the processing of black and white faces. *Psychol Sci,* 15**,** 806-813.

DUNBAR, R. I. M. & SHULTZ, S. 2007. Evolution in the social brain. *Science,* 317**,** 1344-1347.

EDWARDS, J., JACKSON, H. J. & PATTISON, P. E. 2002. Emotion recognition via facial expression and affective prosody in schizophrenia: a methdological review. *Clinical Psychology Review,* 22**,** 789-832.

EISENBERGER, N. I., LIEBERMAN, M. D. & WILLIAMS, K. D. 2003. Does rejection hurt? An fMRI study of social exclusion. *Science,* 302**,** 290-292.

EMERY, N. J. 2000. The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews,* 24**,** 581-604.

FRITH, C. D. 2007. The social brain? *Philos Trans R Soc Lond B,* 362**,** 671-678.

FUNNELL, E. 2001. Evidence for scripts in semantic dementia. Implications for theories of semantic memory. *Cogn Neuropsychol,* 18**,** 323-341.

GALLASE, V. & GOLDMAN, A. 1998. Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences,* 2**,** 493-501.

GOREN, C. C., SARTY, M. & WU, P. Y. 1975. Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics,* 56**,** 544-549.

GUELLAI, B. & STRETI, A. 2011. Cues for early social skills: direct gaze modulates newborns' recognition of talking faces. *PLoS ONE,* 6**,** e18610.

HAMPTON, A. N., BOSSAERTS, P. & O'DOHERTY, J. P. 2006. The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *Journal of Neuroscience,* 26**,** 8360-8367.

HASSON, U., NIR, Y., LEVY, I., FUHRMANN, G. & MALACH, R. 2004. Intersubject synchronization of cortical activity during natural vision. *Science,* 303**,** 1634-1640.

HAXBY, J. V., HOFFMAN, E. A. & GOBBINI, M. I. 2000. The distributed human neural system for face perception. *Trends in Cognitive Sciences,* 4**,** 223-233.

HEIDER, F. 1946. Attitudes and cognitive organization. *Journal of Psychology,* 21**,** 107-112.

HOOKER, C. & PARK, S. 2005. You must be looking at me: the nature of gaze perception in schizophrenia patients. *Cogn Neuropsychiatry,* 10**,** 327-345.

JONES, B. C., DEBRUINE, L. M., LITTLE, A. C., CONWAY, C. A. & FEINBERG, D. R. 2006. Integrating gaze direction and expression in preferences for attractive faces. *Psychol Sci,* 17**,** 588-591.

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.

KAUPPI, J.-P., HUTTUNEN, H., KORKALA, H., JÄÄSKELÄINEN, I. P., SAMS, M. & TOHKA, J. 2011. Face prediction from fMRI data during movie stimulus: strategies for feature selection. *Lecture Notes in Computer Science,* 6792/2011**,** 189-196.

KENNEDY, D. P., GLÄSCHER, J., TYSZKA, J. M. & ADOLPHS, R. 2009. Personal space regulation by the human amygdala. *Nature Neuroscience,* 12**,** 1226-1227.

KINGSTONE, A., FRIESEN, C. K. & GAZZANIGA, M. S. 2000. Reflexive joint attention depends on lateralized cortical connections. *Psychol Sci,* 11**,** 159-166.

LESLIE, A. M. 1987. Pretence and representation in infancy: the origins of theory of mind. *Psychol Rev,* 94**,** 84-106.

LEWIS, P. A., REZAIE, R., BROWN, R., ROBERTS, N. & DUNBAR, R. I. 2011. Ventromedial prefrontal volume predicts understanding of others and social network size. *Neuroimage,* 57**,** 1624-1629.

MARTIN-RODRIGUEZ, J. F. & LEON-CARRION, J. 2010. Theory of mind defifcits in patients with acquired brain injury: a quantitative review. *Neuropsychologia,* 48**,** 1181-1191.

N'DIAVE, K., SANDER, D. & VUILLEUMIER, P. 2009. Self-relevance processing in the human amygdala: gaze direction, facial expression, and emotion intensity. *Emotion,* 9**,** 798-806.

PREMACK, D. & WOODRUFF, G. 1978. Chimpanzee problem-solving: a test for comprhension. *Behavioral and Brain Sciences,* 4**,** 515-526.

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. *Am J Phys Anthropol,* 114**,** 224-241.

TAZUMI, T., HORI, E., MAIOR, R. S., ONO, T. & NISHIJO, H. 2010. Neural correlates to seen gaze-direction and orientation in the macaque monkey amygdala. *Neuroscience,* 169**,** 287-301.

TOMASELLO, M., CALL, J. & HARE, B. 2003. Chimpanzees understand psychological states - the question is which ones and to what extent. *Trends in Cognitive Sciences,* 7**,** 153-156.

VUILLEUMIER, P., ARMONY, J. L., DRIVER, J. & DOLAN, R. J. 2003. Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience,* 6**,** 624-631.

WHALEN, P. J., RAUCH, S. L., ETCOFF, N. L., MCINERNEY, S. C., LEE, M. B. & JENIKE, M. A. 1998. Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience,* 18**,** 411-418.

WINSTON, J. S., STRANGE, B. A., O'DOHERTY, J. & DOLAN, R. J. 2002. Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience,* 5**,** 277-283.

13 Consciousness

It can be said that consciousness is the ultimate and most challenging research question in cognitive neuroscience. Similarly with many other higher cognitive functions, such as attention and emotions, it can be also said that everyone knows what consciousness is. From the perspective of rigorous scientific study consciousness is, however, a highly elusive phenomenon and there are multiple definitions of consciousness. Sometimes it has been even stated that there are as many definitions of consciousness as there are scientists studying consciousness. Consciousness is also a multifaceted phenomenon, and the term consciousness is frequently used to refer to two relatively distinct phenomena.

The first phenomenon that the term consciousness refers to is whether a given person is aware of external stimuli that are being presented to him/her. In the extreme case, of course, this comes down to whether that person is conscious or unconscious (as when in coma or deep sleep). Attempting to answer the intriguing question of at which stage of sensory processing a person becomes aware of the externally applied stimuli has been one of the central themes in neuroscience of consciousness. There are patients who can perceive stimuli without being consciously aware of the stimuli. Consciousness also has a lot in common with attention and working memory and it could be argued that these terms are partially overlapping; each involves limited capacity of, and awareness of the contents in the, limited capacity "processing buffer".

In addition to consciousness meaning awareness of external stimuli, self-awareness is often referred to as ("higher-order") consciousness. It is also characteristic of consciousness that while one is highly cognizant of the contents of one's consciousness (or rather one's self awareness), one is not capable of inspecting introspectively the mechanisms that produce the contents of consciousness. There is a wealth of studies on the neural mechanisms and cognitive prerequisites of self-awareness, including evolutionary ones that have attempted to answer the question of whether (and which) animals are self-aware. Patients with specific types of brain damage can have highly peculiar deficits in consciousness. The so-called "split-brain" patients with severed corpus callosum (*i.e.*, the bundle of white-matter tracts that connects the left and right hemispheres) can exhibit signs suggesting divided consciousness. These lines of research on the neural underpinnings of consciousness will be introduced in this chapter.

13.1 Consciousness as awareness of externally applied stimuli

As stated above, consciousness is often referred to in cognitive neuroscience as awareness of externally applied stimuli. In fact, one is not aware of most of the stimuli that excite one's senses. One good example of this is that even though lower-order sensory neurons react highly specifically to elementary constituents of speech sounds, such as transient sound components that rapidly ascend in sound frequency in the beginning of the consonant-vowel syllable /ba/, one is poorly (if at all) consciously aware of the presence of these sounds. In contrast, one is perfectly aware of the syllables and words that one hears and the semantic meaning those words carry. Based on observations such as these, it has been proposed that consciousness (as per sensory awareness) predominantly involves neural representations of higher-order perceptual objects (Ahissar et al., 2009).

Empirically, the neural mechanisms underlying conscious percepts have been studied with ingenious experimental designs utilizing ambiguous visual (and auditory) stimuli that can be perceived in alternative ways. These are called multi-stable percepts. One example of such stimuli is a picture where there are two faces next to one another. This picture can be perceived either as the two faces or a vase in the middle. In a recent study, electromagnetic activity of human visual cortex was recorded during multi-stable perception of face *vs.* vase (see **Figure 13-1**).

The background (faces) and middle part (vase) were flickered at distinct frequencies, thus allowing tracking of electromagnetic activity that was associated with the percept of face and vase. It was observed that visual cortex activity frequency-followed the percept that was self-reported by the subjects (*via* button presses) as the dominant percept. These results demonstrate that sensory cortex activity is associated with conscious percepts. The authors suggested that these percept-related modulations are caused by top-down influences that enhance the neural representation of the perceived object and thus help segregate objects from the background in the early visual cortex (Parkkonen et al., 2008).

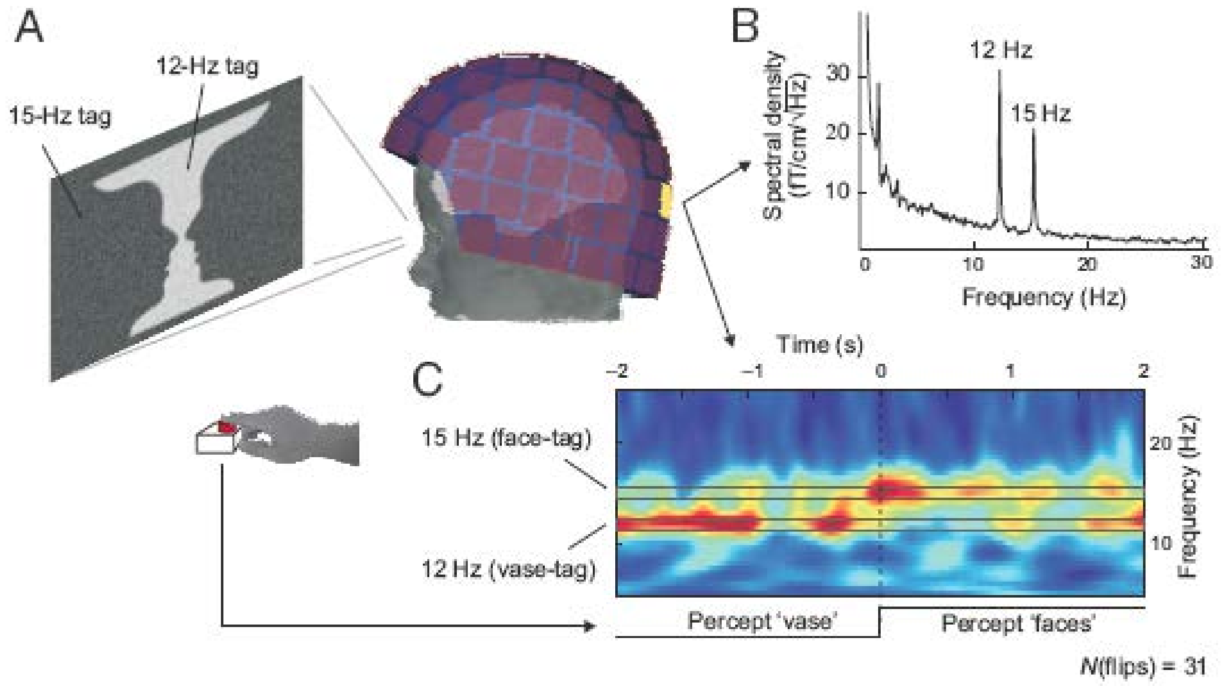


Figure 13-1: Visual-cortical electromagnetic activity correlates with percepts in a multi-stable perception experiment. A) The percept of two dark faces facing one another alternate with the percept of a lighter vase in the middle. Electromagnetic activity is recorded using magnetoencephalography during perception. The sensor that is most sensitive to primary visual cortex activity is located in the back of the sensor array, indicated with yellow color. The subjects indicated with a button press when their percept switched from face to vase and vice versa. Here, the background (*i.e.*, faces) were flickered at 15 Hz, and the center (the vase) was flickered at 12 Hz. B) The frequency-tagging of the face and the vase allowed quantification of distinct 15 and 12 Hz frequency-following responses in the visual cortex recordings. C) As can be seen in time-frequency plots, the power (indicated by the warm colors) increases in visual cortex in the tagged frequency as a function of whether the conscious percept was a face or a vase. Adapted from (Parkkonen et al., 2008).

Oscillatory activity and perceptual binding

In another line of investigation on the neural basis of consciousness, cognitive neuroscientists have carried out studies to solve the so-called perceptual binding problem. The binding problem refers to the puzzle of how different features of a given perceptual object (*e.g.*, color, texture, and shape of an apple) that are processed in parallel in discrete brain areas are ultimately bound together to form a coherent percept. One possible solution that has been suggested to underlie perceptual binding is that neuronal activity synchronizes in specific frequency bands across the involved brain areas to enable coherent network representations to emerge (Gray et al., 1989). Supporting these ideas, there are findings indicating that gamma-band activity, recorded non-invasively with electroencephalography, is enhanced in subjects trained to perceive a Dalmatian dog hidden in noise compared with the responses of naïve subjects (Bertrand and Tallon-Baudry, 2000). These results are summarized in **Figure 13-2** below.

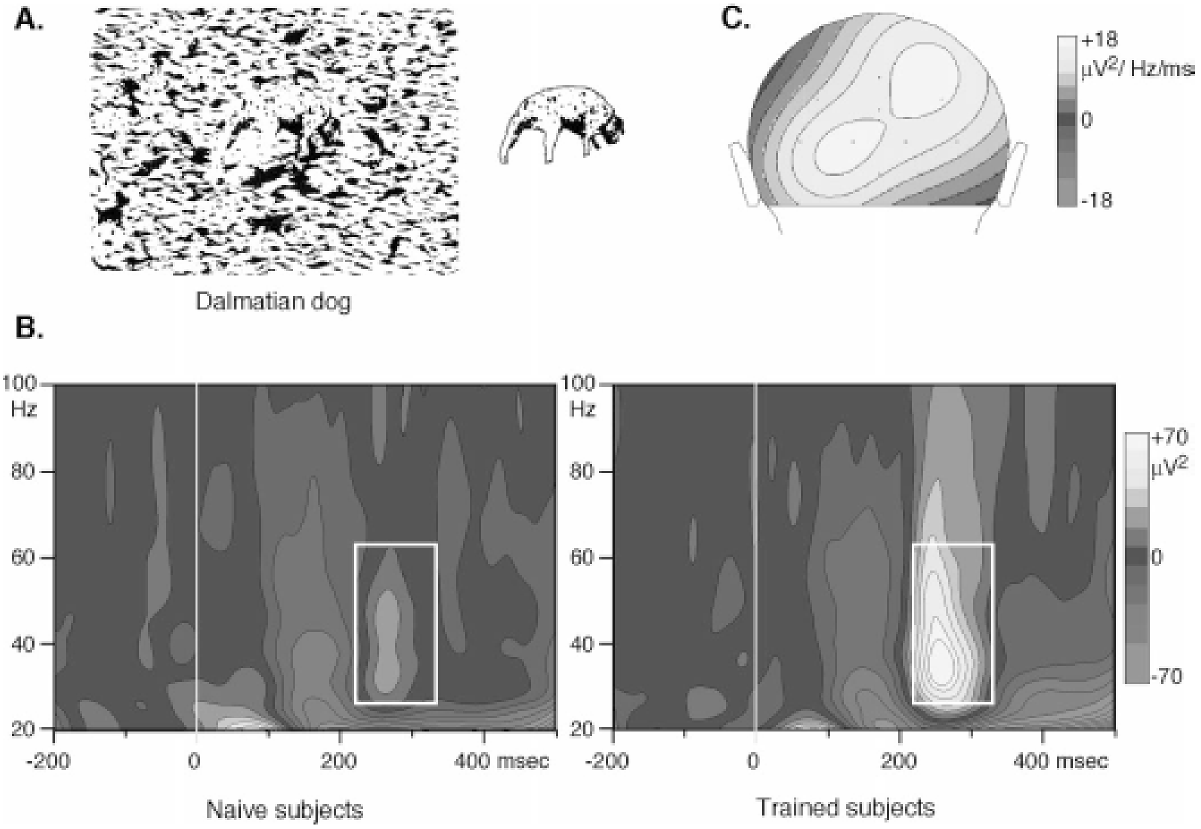


Figure 13-2. Perception of a Dalmatian dog in noise correlates with emergence of gamma-band synchronized activity in the electroencephalogram of healthy subjects. A) The Dalmatian hidden in noise; it is initially difficult to perceive the dog, but once it is indicated to subjects, perception is relatively automatic. B) Time-frequency representations (where lighter grayscale indicates enhancement) of stimulus-induced electroencephalogram activity shows emergence of a gamma frequency band response in subjects trained to detect the Dalmatian as compared with naïve subjects. C) Shown is the scalp distribution of the electroencephalogram gamma-band responses that suggests visual cortical source activity (adapted from (Bertrand and Tallon-Baudry, 2000)).

Consciousness as an emergent phenomenon

In keeping with the hypothesis that synchronization of neural activity across brain areas underlies perceptual binding, consciousness is currently viewed to be an emergent phenomenon that involves networks of brain areas instead a single "consciousness area", even though certain brain areas, including prefrontal cortex and anterior cingulate cortex, might be more central for consciousness than others (Dehaene and Naccache, 2001). Further, the precise constellation of which brain areas form the network that gives rise to conscious experience at a given instant is thought to depend on the concurrent contents of consciousness that is regulated by attention (and neuroanatomically by the brain regions that underlie attentional selection including prefrontal, anterior cingulate, and posterior parietal cortices). For instance, consciousness during auditory stimulation, such as when fully immersed in listening to a musical performance by a symphony orchestra, probably involves a network of brain areas that is at least to some extent different from the brain areas involved in conscious experience during intensive visual stimulation, such as when playing a video game involving fast car-racing on a track.

Perception without conscious awareness

Perception of external stimuli without conscious awareness has been demonstrated in patients who have lost sight due to damage to their visual system that has spared subcortical visual connections to the superior colliculi. These patients are effectively blind in that they are not aware of being able to perceive any visual stimuli. However, presenting spots of light to left *vs.* right hemifield and prompting the patients to guess the side at which the stimuli were presented results in well above chance level guessing. This phenomenon has been termed blindsight (Weiskrantz, 1990). Perception of emotional facial expressions by amygdala without conscious awareness when very brief stimulus presentation times (ca. 25 ms) are used (as was described in Chapter 12) provide another indication that perception can occur without conscious awareness (Whalen et al., 1998). These findings, showing primitive perception by subcortical structures without conscious experience of the stimuli, have been taken to suggest that conscious perception is something that is unique to cerebral cortex.

13.2 Consciousness as self-awareness

Self-awareness is an intriguing phenomenon; in addition to being aware of oneself as an active agent with one's own thoughts and intentions, being self-aware also implies having a continuous percept of self. Indeed, despite that one's memories, skills, things one likes and dislikes, and even to some extent one's personality, change in the course of life, one maintains a fairly stable self-percept. One approach that has been utilized in the study of neural and cognitive prerequisites of self-awareness is investigation of what other species in addition to humans are capable of self-awareness. Cross-species comparisons make it possible to answer the question of which neural structures and cognitive capabilities correlate with the capability for self-awareness. These findings will be described in the following.

Evolution of self-awareness

There are a number of studies that have attempted to test whether other species besides humans exhibit signs of self-awareness. In animal studies, it is of course not possible to ask directly whether the animal is self-aware, but rather it has to be indirectly inferred from the behavior of the animal. One widely used test is the test for self-recognition of one's own mirror image. In one early study, a group of chimpanzees were placed in front of a mirror for several days. After this period of familiarizing, a mark was applied above the eye of each animal while the animal was under anesthesia. Upon awakening, all of the tested animals started touching the mark when seeing it in the mirror, thus suggesting that chimpanzees do have self-awareness (see **Figure 13-3**). Monkeys tested with the same procedure failed to exhibit signs of self-recognition, and thus signs of self-awareness (Gordon, 1970). Out of a number of different species tested, chimpanzees, orangutans, dolphins (in modified tests), and to some extent gorillas have been the only ones to pass the mirror self-recognition test (for a review, see (Keenan et al., 2005)).

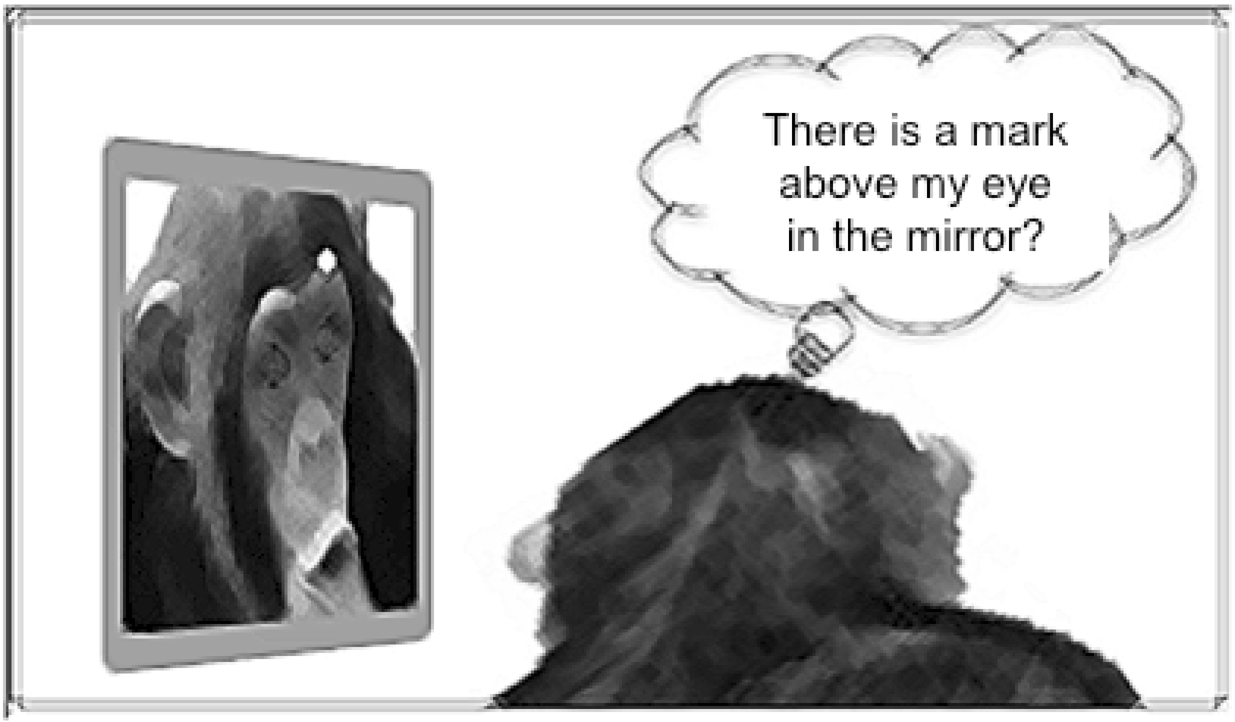


Figure 13-3: Chimpanzees pass the mirror self-recognition test by touching their own face to explore a mark above their eyebrow when they see it in their own reflection in the mirror. This has been taken to suggest that chimpanzees are self-aware. See text for details.

On what grounds is it claimed that the mirror self-recognition test reflects self-awareness? One approach that has been used to validate this method are studies in infants and young children where the emergence of mirror self-recognition ability has been correlated with the emergence of other measures of self-awareness, for example, self-conscious emotions such as feelings of shame (Lewis, 1992). The use of personal pronouns also begins at around the age that children pass the mirror self-recognition test (Imbens-Bailey and Pan, 1998). It has further been observed that species passing the mirror self-recognition test also pass theory of mind tests; it can be argued that self-awareness is needed in order to be able to infer the intentions of others (Keenan et al., 2005). In children, the ability to deceive (which is an indication of development of theory of mind) also begins to emerge at around the age that self-awareness begins to form (Ritblatt, 2000).

Neural basis of self-awareness: insights from split-brain patients

Investigating the neural basis of self-awareness is a highly challenging task. There are, however, patients with a specific type of brain damage that have allowed scientists to obtain highly interesting findings that shed light on the neural basis of self-awareness. These so-called split-brain patients were treated for drug-resistant epilepsy by surgically cutting their inter-hemispheric connections. Depending on the patient either major parts or all of corpus callosum were cut, and in some cases also the anterior and posterior commissura. In these patients it has been possible to study the functions of the two hemispheres when they are isolated from one another, and indeed a wealth of interesting observations have been obtained (Gazzaniga, 2000). What makes the findings in split-brain patients especially interesting from the point of view of consciousness is that there is evidence suggesting that consciousness is to some extent split in the split-brain patients.

An example of empirical studies that have been carried out in split-brain patients is illustrated in **Figure 13-4**. Information presented to the right hemifield (in this case the name of the object to be picked up from behind the screen) is processed in the left hemisphere that controls the right hand. Since the corpus callosum is severed, information of the object name does not pass to the right hemisphere and the patient cannot retrieve the desired object with his/her left hand. However, with right hand that is controlled by the left hemisphere, the object is retrieved without effort. This design can be reversed, pending that the verbal instructions directed to the language non-dominant right hemisphere are kept fairly simple and concrete. Thus, the two hemispheres seem to be conscious of separate things at least under the experimental conditions that were used.

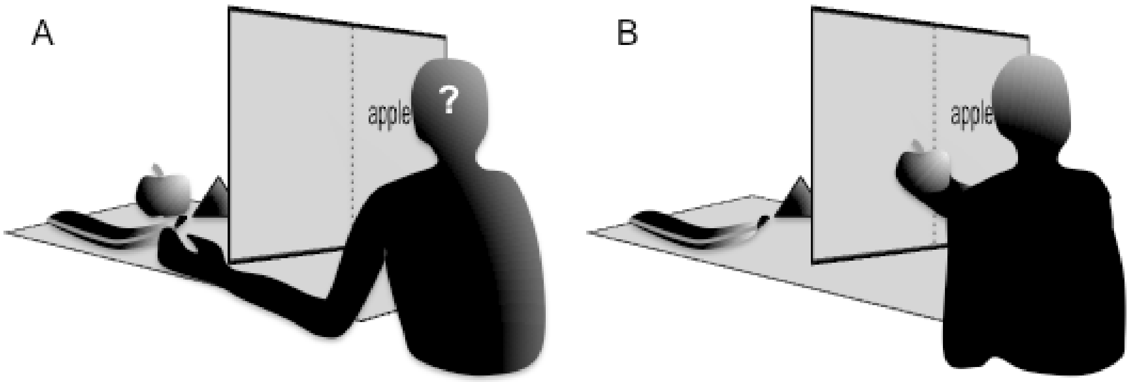


Figure 13-4: An example of the type of tests that have been carried out in split-brain patients. A) Here, the name of object ("apple") to be selected by the left hand of the patient from behind the screen is presented to the right visual field where the visual information travels to the left hemisphere. Since the information cannot pass *via* the corpus callosum to the right hemisphere that controls the left hand, the patient fails to pick the right object. B) When the patient is allowed to use his/her right hand, controlled by the left hemisphere, the task is accomplished with ease.

The extent that there is self-awareness in the right hemisphere has been debated. It has been suggested, based on both self-reports and experimental manipulations in split brain patients, that the left hemisphere is the one that interprets information relevant to self and makes it possible to maintain coherent self-percept, which is one of the key aspects of higher-order consciousness (Gazzaniga, 2000). The patients also themselves verbally report that their self-percept and consciousness are both entirely intact, which has been taken as further evidence pointing to the left hemisphere as the site of higher-order consciousness.

This line of reasoning has been supported by claims according to which highly developed language functions are a prerequisite for consciousness. On the other hand, in their everyday lives, split-brain patients have been reported to have conflicting motives and intentions. For example, the left hand of a split-brain patient might attempt to do something improper that the right hand then prevents from happening. Observations such as these suggest that consciousness is to some extent split in these patients.

Signs of right hemisphere self-awareness in split brain patients

It is possible that the fact that the right hemisphere cannot give a verbal account of the contents of its awareness has sometimes been interpreted as a sign indicating that the left speech-dominant hemisphere is the site of the true consciousness in these patients and that right-hemisphere consciousness would be lesser in nature. Interestingly, in a version of the mirror self-recognition test it is the right hemisphere that appears to yield more robust response than the left hemisphere. Specifically, measuring skin-conductance responses of two split brain patients to familiar *vs.* own face presented to left *vs.* right hemispheres, it was observed that response to own face compared with familiar faces was much larger when presented to the right hemisphere than to the left hemisphere (Preilowski, 1977). Thus, it appears that at least on certain measures, the right-hemisphere self-awareness in split-brain patients is more robust than what has been thought (Keenan et al., 2005).

Episodic memory also appears to be intact and even more accurate in the right than the left hemisphere (Phelps and Gazzaniga, 1992), which suggests that some type of self-perception is accomplished also by the right hemisphere. The nature of right-hemisphere consciousness / self-awareness is, however, still largely unexplored, mainly due to difficulties in obtaining reports from the non-verbal right hemisphere of split-brain patients. The current understanding in cognitive neuroscience is that the left hemisphere has many processing abilities that support higher-order consciousness and self-awareness (Gazzaniga, 2000). The non-verbal right hemisphere does, however, appear to be self-aware, shows motivated behavior, and is on some perceptual processing tasks superior to the left hemisphere. Naturally, in healthy subjects the different aspects of self-awareness/consciousness of the left and right hemispheres are combined.

Consciousness and subconscious influences

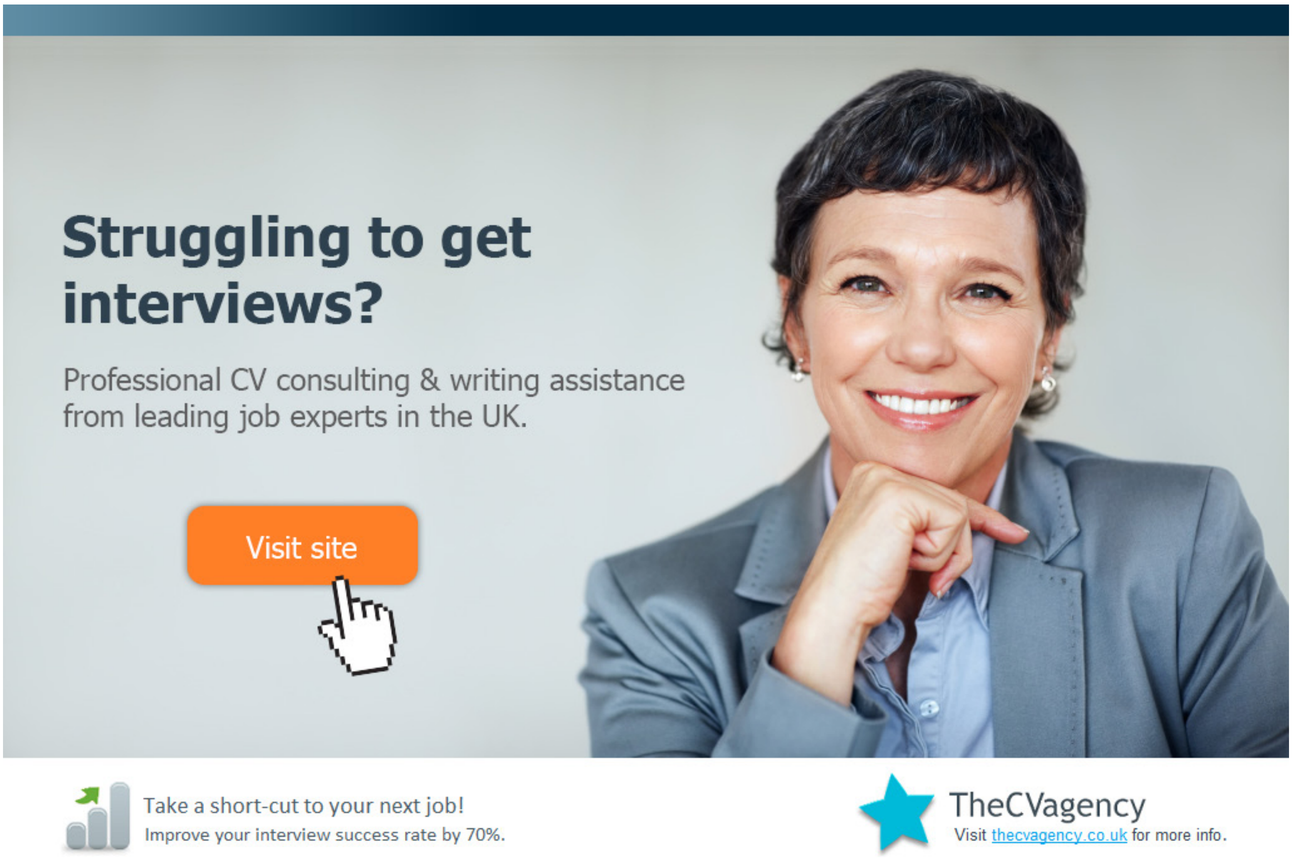
The limited capacity of consciousness is a central theme in the—at it's time pioneering—theoretical work of Sigmund Freund on human personality. The theory Freud developed has become widely known as the psychoanalytic theory of personality. In this theory, the human mind is seen as consisting of what Freud termed id, ego, and superego (Freud, 1923). The id refers to subconscious impulses and motives (in the theory typically sexual) that constantly urge one to take actions. Superego is the counterforce, comprising of learned constraints to behavior and what is appropriate behavior given the social context. In the psychoanalytic theory, ego equates with the limited-capacity self-awareness. While the psychonanalytic theory of personality has been challenged on multiple grounds, it cannot be denied that it represented a highly significant step forward at the time that Freud proposed it in the turn of the 19th and 20th Centuries. The realization that subconscious impulses play a role in influencing behavior is especially notable.

Central executive's limited capacity may explain subconscious influences

In modern cognitive neuroscience terms the subconscious impulses can perhaps be best described as the limited capacity of attention and working memory making it impossible for one to be aware of all the factors that influence one's behavior. Furthermore, schemas that help one to perceive the world in a meaningful and coherent manner are based on tens, hundreds, or even thousands of encounters with particular types of situations and persons, thus making it impossible to be aware of all the factors that have shaped the schemas (and thus one's perception of the world). Rather, one operates using the higher-order object and event schemata that are generalizations of the world, without being aware of the instances based on which the schemata were formed.

Lack of generalized knowledge of the world is manifested in the behavior of small children. I once wondered about the behavior of my, at the time approximately one-and-a-half year old, daughter when we were sitting outside on our patio and eating fresh strawberries. Instead of eating each strawberry before reaching for another one, she was biting a piece off of each one and immediately grabbing the next one for a single bite. I wondered that possibly to her little eyes the strawberries were not as identical as for an adult who is looking at them and activating the generalized object representation for a strawberry. Rather, the individual strawberries probably seemed to her to be of different shapes, slightly different colors and tastes, and thus she wanted to try and taste as many as she could. On the other hand, through such sampling, children ultimately develop the highly generalized object representations that then make it possible to generalize, for instance, a wealth of different kinds of chairs as something that one can sit on.

The ability to generalize and operate at the level of higher-order representations gives the benefit of being able to make quick decisions. In fact, if one would have to be aware of all the factors that influence one's decision making, even the smallest of tasks would take infinitely long. The implicit knowledge that Freud termed as the subconscious also relates to the phenomenon of intuitive decision-making—the first impression or a "gut feeling" on what is the best option when faced with a complex decision where there are multiple options without a salient best choice. While intuitive decision-making can be biased compared with analytical decision-making, trusting intuition has been noted to have its advantages (Hogarth, 2005).



13.3 Concluding remarks

Even though consciousness is a highly elusive phenomenon with multiple definitions, systematic research on consciousness has delineated certain basic principles that different aspects of conscious experience, namely awareness of external stimuli, awareness of the self and perception of continuity of the self through the lifespan of an individual, are grounded on. First, it appears that one is aware of sensory stimuli due to activation of a network of brain areas that underlie higher-order perceptual object and event representations, and that synchronization of brain activity across distinct brain areas helps bind the different aspects of the conscious percept together. Furthermore, conscious experience seems to be a quality that is restricted to cerebral cortex, with brain areas underlying attentional selection, including prefrontal, anterior cingulate, and posterior parietal cortex, forming the "core" of dynamic networks of brain areas that give rise to consciousness.

Self-awareness that has been also termed higher-order consciousness does not seem to be entirely unique to humans, as certain other species including chimpanzees and dolphins seem to be self-aware. Studies on split-brain patients have shown that the language-dominant left hemisphere seems to be especially important for building the self-narrative and coherent percept of the self while going through one's life. Some aspects of self-awareness do, however, seem to reside also in the right hemisphere of split-brain patients. Overall, while several answers have been provided by cognitive neuroscience studies of consciousness, the grand puzzle of how the human consciousness emerges still remains to be solved in future research.

13.4 References

AHISSAR, M., NAHUM, M., NELKEN, I. & HOCHSTEIN, S. 2009. Reverse hierarchies and sensory learning. *Philos Trans R Soc Lond B Biol Sci,* 364**,** 285-299.

BERTRAND, O. & TALLON-BAUDRY, C. 2000. Oscillatory gamma activity in humans: a possible role for object representation. *International Journal of Psychophysiology,* 38**,** 211-223.

DEHAENE, S. & NACCACHE, L. 2001. Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition,* 79**,** 1-37.

FREUD, S. 1923. *The ego and the id,* London, Hogarth Press.

GAZZANIGA, M. S. 2000. Cerebral specialization and interhemispheric communication: does the corpus callosum enable the human condition. *Brain,* 123**,** 1293-1326.

GORDON, G. G. J. 1970. Chimpanzees: self-recognition. *Science,* 167**,** 86-87.

GRAY, C. M., KÖNIG, P., ENGEL, A. K. & SINGER, W. 1989. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature,* 338**,** 334-337.

HOGARTH, R. M. 2005. Deciding analytically or trusting your intuition? The advantages and disadvantages of analytic and intuitive thought. *In:* TILMANN, B. & HABERSTROH, H. (eds.) *The routines of decision making.* New Jersey: Lawrence Erlbaum Associates.

IMBENS-BAILEY, A. & PAN, A. 1998. The pragmatics of self- and other reference in young children. *Social Development,* 7**,** 219-233.

KEENAN, J. P., RUBIO, J., RACIOPPI, C., JOHNSON, A. & BARNACZ, A. 2005. The right hemisphere and the dark side of consciousness. *Cortex,* 41**,** 695-704.

LEWIS, M. 1992. Self-conscious emotions and the development of self. *In:* SHAPIRO, T. & EMDE, R. N. (eds.) *Affect: psychoanalytic perspectives.* Madison, CT: International Universities Press.

PARKKONEN, L., ANDERSSON, J., HÄMÄLÄINEN, M. S. & HARI, R. 2008. Early visual brain areas reflect the percept of an ambiguous scene. *Proc Natl Acad Sci USA,* 105**,** 20500-20504.

PHELPS, E. A. & GAZZANIGA, M. S. 1992. Hemispheric differences in mnemonic processing: the effects of left hemisphere interpretation. *Neuropsychologia,* 30**,** 293-297.

PREILOWSKI, B. 1977. Self-recognition as a test of consciousness in the left and right hemisphere of "split-brain" patients. *Activitas Nervosa Superior (Praha),* 19 (Suppl 2)**,** 343-344.

RITBLATT, S. N. 2000. Children's level of participation in a false-belief task, age, and theory of mind. *Journal of Genetic Psychology,* 161**,** 53-64.

WEISKRANTZ, L. 1990. *Blindsight: a case study and implications,* Oxford, Oxford University Press.

WHALEN, P. J., RAUCH, S. L., ETCOFF, N. L., MCINERNEY, S. C., LEE, M. B. & JENIKE, M. A. 1998. Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience,* 18**,** 411-418.